

## THE GREAT CRASH, THE OIL PRICE SHOCK, AND THE UNIT ROOT HYPOTHESIS

BY PIERRE PERRON<sup>1</sup>

We consider the null hypothesis that a time series has a unit root with possibly nonzero drift against the alternative that the process is “trend-stationary.” The interest is that we allow under both the null and alternative hypotheses for the presence of a one-time change in the level or in the slope of the trend function. We show how standard tests of the unit root hypothesis against trend stationary alternatives cannot reject the unit root hypothesis if the true data generating mechanism is that of stationary fluctuations around a trend function which contains a one-time break. This holds even asymptotically. We derive test statistics which allow us to distinguish the two hypotheses when a break is present. Their limiting distribution is established and selected percentage points are tabulated. We apply these tests to the Nelson-Plosser data set and to the postwar quarterly real GNP series. In the former, the break is due to the 1929 crash and takes the form of a sudden change in the level of the series. For 11 out of the 14 series analyzed by Nelson and Plosser we can reject at a high confidence level the unit root hypothesis. In the case of the postwar quarterly real GNP series, the break in the trend function occurs at the time of the oil price shock (1973) and takes the form of a change in the slope. Here again we can reject the null hypothesis of a unit root. If one is ready to postulate that the 1929 crash and the slowdown in growth after 1973 are not realizations of an underlying time-invariant stochastic process but can be modeled as exogenous, then the conclusion is that most macroeconomic time series are not characterized by the presence of a unit root. Fluctuations are indeed stationary around a deterministic trend function. The only “shocks” which have had persistent effects are the 1929 crash and the 1973 oil price shock.

**KEYWORDS:** Hypothesis testing, intervention analysis, structural change, stochastic trends, deterministic trends, functional weak convergence, Wiener process, macroeconomic time series.

### 1. INTRODUCTION

THE UNIT ROOT HYPOTHESIS has recently attracted a considerable amount of work in both the economics and statistics literature. Indeed, the view that most economic time series are characterized by a stochastic rather than deterministic nonstationarity has become prevalent. The seminal study of Nelson and Plosser (1982) which found that most macroeconomic variables have a univariate time series structure with a unit root has catalyzed a burgeoning research program with both empirical and theoretical dimensions.

Nelson and Plosser's study was followed by a series of empirical analyses which basically confirmed their findings. Some (Stulz and Wasserfallen (1985) and Wasserfallen (1986), among others) applied a similar Dickey–Fuller (1979) statistical methodology to other economic series. On the statistical front, there

<sup>1</sup>I wish to thank Brian Campbell, Larry Christiano, Jean-Marie Dufour, Clive Granger, Whitney Newey, Hashem Pesaran, the referees, and the editor for useful comments. Christian Dea and Nicholas Marceau provided useful research assistance. This research was supported by the Social Sciences and Humanities Research Council of Canada, the Natural Sciences and Engineering Council of Canada, and Québec's F.C.A.R. grants. The first draft of this paper was written while the author was Assistant Professor at the Université de Montréal.

emerged an interest in developing alternative approaches to test the unit root hypothesis. Examples include: the class of tests proposed by Phillips and Perron (1988) and the methodology suggested by Campbell and Mankiw (1987, 1988) and Cochrane (1988) using an estimate of the spectral density at frequency zero. Empirical applications of these methodologies generally reaffirmed the conclusion that most macroeconomic time series have a unit root (e.g., Perron (1988)).

These studies had many effects on economic theorizing. They seem to confirm previous analyses which had advanced the unit root hypothesis for particular economic series, e.g., consumption (Hall (1978)), velocity of money (Gould and Nelson (1974)), and stock prices (Samuelson (1973)). They also launched a series of theoretical investigations with implications consistent with a unit root, e.g., Blanchard and Summers (1986) for employment. Furthermore, a considerable stock of statistical tools was developed for more general models with integrated variables; these include the cointegration framework (Engle and Granger (1987)) and multivariate systems (Stock and Watson (1988) and Phillips and Durlauf (1986)).

As far as macroeconomic theories are concerned, the most important implication of the unit root revolution, is that under this hypothesis random shocks have a permanent effect on the system. Fluctuations are not transitory. This implication, as forcefully argued by Nelson and Plosser, has profound consequences for business cycle theories. It runs counter to the prevailing view that business cycles are transitory fluctuations around a more or less stable trend path. It is therefore of importance to assess carefully the reliability of the unit root hypothesis as an empirical fact.

The aim of this paper appears startling, given the results in the above mentioned literature. Our conclusion is that most macroeconomic time series are not characterized by the presence of a unit root and that fluctuations are indeed transitory. Only two events (shocks) have had a permanent effect on the various macroeconomic variables: the Great Crash of 1929 and the oil price shock of 1973.

Of course, to reach such a conclusion, a particular postulate must be introduced which differentiates our approach from the previous ones. This postulate is that the Great Crash and the oil price shock were not a realization of the underlying data-generating mechanism of the various series. In this sense, we consider these shocks as exogenous. The exogeneity assumption is not a statement about a descriptive model for the time series representation of the variables. It is used here as a device to remove the influence of these shocks from the noise function. A more detailed discussion of these issues and their implications can be found in Section 6.

These two shocks are rather different in nature. On one hand, the Great Crash created a dramatic drop in the mean of most aggregate variables. On the other hand, the 1973 oil price shock was followed by a change in the slope of the trend for most aggregates, i.e., a slowdown in growth. In this light, our aim is to show that most macroeconomic variables are "trend-stationary" if one allows a single

change in the intercept of the trend function after 1929 and a single change in the slope of the trend function after 1973.

Our approach is in the spirit of the "intervention analysis" suggested by Box and Tiao (1975). According to their methodology, "aberrant" or "outlying" events can be separated from the noise function and be modeled as changes or "interventions" in the deterministic part of the general time series model. Using such a strategy makes it "possible to distinguish between what can and what cannot be explained by the noise" (Box and Tiao (1975, p. 72)). These "interventions" are assumed to occur at a known date. The same strategy is used in the present analysis in that we consider the time of the changes in the trend function as fixed rather than as a random variable to be estimated.

To make our point as unambiguous as possible, we use the same data set as Nelson and Plosser, as well as the real GNP series analyzed by Campbell and Mankiw. The data set used by Nelson and Plosser contains fourteen macroeconomic variables sampled annually. All series end in 1970 and contain only one break, the 1929 Great Crash. We shall not analyze the unemployment rate series since there is a general agreement that it is stationary. The real GNP series is postwar quarterly from 1947:I to 1986:III and so contains only one break as well, the 1973 oil shock. Furthermore, to make our analysis as similar as possible to previous ones, the statistical methodology applied here is an extension of the Dickey-Fuller methodology (as used by Nelson and Plosser) to test for the presence of a single unit root in a univariate time series.

The plan of the paper is as follows. Section 2 motivates the ensuing analysis and presents the alternative models considered. Section 3 shows that usual tests will not be able to reject the unit root hypothesis if in fact the deterministic trend of the series has a single break (either in the intercept or the slope). In Section 4, we develop formal statistical tests of the null hypothesis of a unit root which can distinguish the unit root hypothesis from that of a stationary series around a trend which has a single break. The asymptotic distributions under the null hypothesis are derived and tabulated. Empirical results from applying these procedures are presented in Section 5. Section 6 contains a discussion of some issues raised by our analysis and suggestions for future research. All theorems are proved in Appendix A.

## 2. MOTIVATION

The null hypothesis considered is that a given series  $\{y_t\}_0^T$  (of which a sample of size  $T + 1$  is available) is a realization of a time series process characterized by the presence of a unit root and possibly a nonzero drift. However, the approach is generalized to allow a one-time change in the structure occurring at a time  $T_B$  ( $1 < T_B < T$ ). Three different models are considered under the null hypothesis: one that permits an exogeneous change in the level of the series (a "crash"), one that permits an exogenous change in the rate of growth, and one that allows both

change. These hypotheses are parameterized as follows:

*Null hypotheses:*

$$\text{Model (A)} \quad y_t = \mu + dD(TB)_t + y_{t-1} + e_t,$$

$$\text{Model (B)} \quad y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1)DU_t + e_t,$$

$$\text{Model (C)} \quad y_t = \mu_1 + y_{t-1} + dD(TB)_t + (\mu_2 - \mu_1)DU_t + e_t, \quad \text{where}$$

$$D(TB)_t = 1 \quad \text{if } t = T_B + 1, \quad 0 \text{ otherwise;}$$

$$DU_t = 1 \quad \text{if } t > T_B, \quad 0 \text{ otherwise; and}$$

$$A(L)e_t = B(L)v_t,$$

$v_t \sim \text{i.i.d. } (0, \sigma^2)$ , with  $A(L)$  and  $B(L)$   $p$ th and  $q$ th order polynomials, respectively, in the lag operator  $L$ .

The innovation series  $\{e_t\}$  is taken to be of the ARMA( $p, q$ ) type with the orders  $p$  and  $q$  possibly unknown. This postulate allows the series  $\{y_t\}$  to represent quite general processes. More general conditions are possible and will be used in subsequent theoretical derivations.

Instead of considering the alternative hypothesis that  $y_t$  is a stationary series around a deterministic linear trend with time invariant parameters, we shall analyze the following three possible alternative models:

*Alternative hypotheses:*

$$\text{Model (A)} \quad y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)DU_t + e_t,$$

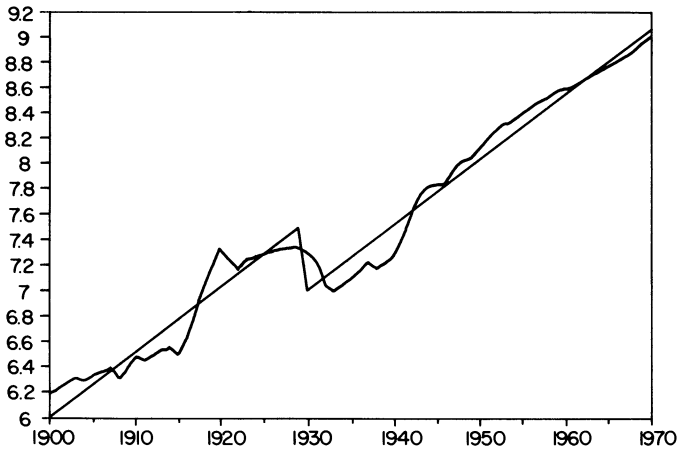
$$\text{Model (B)} \quad y_t = \mu + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + e_t,$$

$$\text{Model (C)} \quad y_t = \mu_1 + \beta_1 t + (\mu_2 - \mu_1)DU_t + (\beta_2 - \beta_1)DT_t + e_t,$$

where

$$DT_t^* = t - T_B, \quad \text{and} \quad DT_t = t \quad \text{if } t > T_B \quad \text{and } 0 \text{ otherwise.}$$

Here,  $T_B$  refers to the time of break, i.e., the period at which the change in the parameters of the trend function occurs. Model (A) describes what we shall refer to as the crash model. The null hypothesis of a unit root is characterized by a dummy variable which takes the value one at the time of break. Under the alternative hypothesis of a "trend-stationary" system, Model (A) allows for a one-time change in the intercept of the trend function. For the empirical cases we have in mind,  $T_B$  is the year 1929 and  $\mu_2 < \mu_1$ . Model (B) is referred to as the "changing growth" model. Under the alternative hypothesis, a change in the slope of the trend function without any sudden change in the level at the time of the break is allowed. Under the null hypothesis, the model specifies that the drift parameter  $\mu$  changes from  $\mu_1$  to  $\mu_2$  at time  $T_B$ . In the empirical examples presented in Section 5,  $T_B$  is the first quarter of 1973 and  $\beta_2 < \beta_1$ , reflecting a slowdown in growth following the oil shock. Model (C) allows for both effects to take place simultaneously, i.e., a sudden change in the level followed by a different growth path.



Note: The broken straight line is a fitted trend (by OLS) of the form  $\tilde{y}_t = \tilde{\mu} + \tilde{\gamma} DU_t + \tilde{\beta} t$  where  $DU_t = 0$  if  $t \leq 1929$  and  $DU_t = 1$  if  $t > 1929$ .

FIGURE 1.—Logarithm of “Nominal Wages.”

To motivate the use of these three models as possible alternatives to the unit root with drift hypothesis, we present in this section some descriptive analyses for three series: “nominal wages” (1900–1970), “quarterly real GNP” (1947:I–1986:III) and “common stock prices” (1871–1970).

Figure 1 shows a plot of the logarithm of the nominal wage series. A feature of this graph is the marked decrease between 1929 and 1930. Apart from this change, the trend appears fairly stable (same slope) over the entire period. The solid line is the estimated trend line from a regression on a constant, a trend and a dummy variable taking a value of 0 prior and at 1929 and value 1 afterwards. Table I presents the results from estimating (by OLS) a regression of the Dickey-Fuller type, i.e.:

$$(1) \quad y_t = \tilde{\mu} + \tilde{\beta} t + \tilde{\alpha} y_{t-1} + \sum_{i=1}^k \tilde{c}_i \Delta y_{t-i} + \tilde{\epsilon}_t.$$

The first row presents the full sample regression. The coefficient on the lag dependent variable is 0.910 with a  $t$  statistic for the hypothesis that  $\alpha = 1$  of  $-2.09$ . Using the critical values tabulated by Dickey and Fuller, we cannot reject the null hypothesis of a unit root. When the sample is split in two (pre-1929 and post-1929), the estimated value of  $\alpha$  decreases dramatically: 0.304 for the pre-1929 sample and 0.735 for the post-1929 sample. However, due to the small samples available, the  $t$  statistics are not large enough (in absolute value) to reject the hypothesis that  $\alpha = 1$ , even at the 10 percent level.

Two features are worth emphasizing from this example: (a) the full sample estimate of  $\alpha$  is markedly superior to any of the split sample estimates and relatively close to one. It appears that the 1929 crash is responsible for the near unit root value of  $\alpha$ ; and (b) the split sample regressions are not powerful enough

TABLE I  
REGRESSION ANALYSIS FOR THE WAGES, QUARTERLY GNP, AND COMMON STOCK PRICE SERIES

Regression: $y_t = \bar{\mu} + \bar{\beta}t + \bar{\alpha}y_{t-1} + \sum_{i=1}^k \bar{c}_i \Delta y_{t-i} + \bar{\varepsilon}_t$								
Series/Period	$k$	$\bar{\mu}$	$t_{\bar{\mu}}$	$\bar{\beta}$	$t_{\bar{\beta}}$	$\bar{\alpha}$	$t_{\bar{\alpha}}$	$S(\bar{\varepsilon})$
(a) Wages								
1900–1970 <sup>a</sup>	2	0.566	2.30	0.004	2.30	0.910	-2.09	0.060
1900–1929	7	4.299	2.84	0.037	2.73	0.304	-2.82	0.0803
1930–1970	8	1.632	3.60	0.012	2.64	0.735	-3.19	0.0269
(b) Common stock prices								
1871–1970 <sup>a</sup>	2	0.481	2.02	0.003	2.37	0.913	-2.05	0.158
1871–1929	3	0.3468	2.13	0.0063	2.70	0.732	-2.29	0.1209
1930–1970	4	-0.5312	-1.64	0.0166	1.96	0.788	-1.89	0.1376
(c) Quarterly real GNP								
1947:I–1986:III	2	0.386	2.90	0.0004	2.71	0.946	-2.85	0.010
1947:I–1973:I	2	0.637	3.04	0.0008	2.99	0.910	-3.02	0.0099
1973:II–1986:III	1	0.883	2.23	0.0008	2.27	0.878	-2.23	0.0102

<sup>a</sup>Results taken from Nelson and Plosser (1982, Table 5).

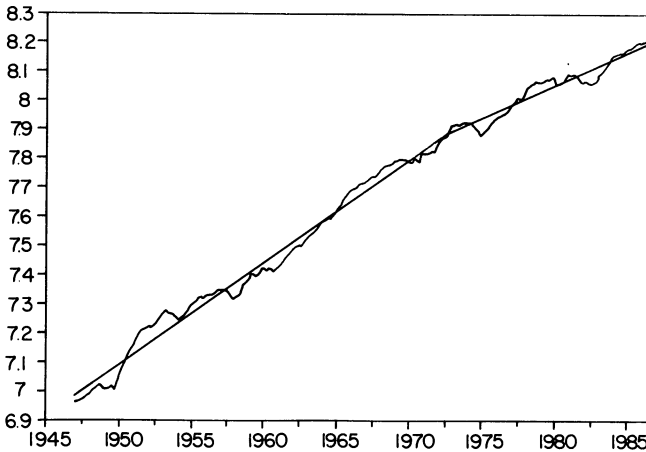
to reject the hypothesis that  $\alpha = 1$  even though the estimates are well below one. It would be useful, in this light, to have a more powerful procedure based on the full sample that would allow the 1929 break to be exogenous.

Figure 2 graphs the postwar quarterly real GNP series. Here, the series behave according to Model (B) where there is no sharp change in the level of the series at the 1973:I break point but rather a change in the slope. The solid line is a fitted trend where a dummy variable is included in the regression, taking the value 0 prior and at 1973:I and the value  $(t - 105)$  after 1973:I (1973:I being the 105th observation in the sample). Table I compares regressions of the form (1) with full and split samples. Again, the estimate of  $\alpha$  is lower in both subsamples than with

TABLE II  
SAMPLE AUTOCORRELATIONS OF THE "DETRENDED" SERIES

Series	Period	$T$	Variance	$r_1$	$r_2$	$r_3$	$r_4$	$r_5$	$r_6$
Real GNP	A 1909–1970	62	0.010	0.77	0.45	0.23	0.11	0.05	0.04
Nominal GNP	A 1909–1970	62	0.023	0.68	0.31	0.12	0.08	0.11	0.12
Real per capita GNP	A 1909–1970	62	0.012	0.81	0.54	0.33	0.20	0.13	0.09
Industrial production	A 1860–1970	111	0.017	0.71	0.44	0.32	0.17	0.08	0.12
Employment	A 1890–1970	81	0.005	0.82	0.59	0.43	0.30	0.20	0.15
GNP deflator	A 1889–1970	82	0.015	0.82	0.63	0.45	0.31	0.17	0.06
Consumer prices	A 1860–1970	111	0.066	0.96	0.89	0.80	0.71	0.63	0.54
Wages	A 1900–1970	71	0.016	0.76	0.47	0.26	0.12	0.03	-0.03
Real wages	C 1900–1970	71	0.003	0.74	0.40	0.12	-0.12	-0.27	-0.33
Money stock	A 1889–1970	82	0.023	0.87	0.69	0.52	0.38	0.25	0.11
Velocity	A 1860–1970	102	0.036	0.90	0.79	0.70	0.62	0.57	0.52
Interest rate	A 1900–1970	71	0.587	0.77	0.58	0.38	0.25	0.15	0.11
Common stock prices	C 1871–1970	100	0.066	0.80	0.53	0.36	0.20	0.10	0.08
Quarterly GNP	B 47:I 86:III	159	0.001	0.94	0.83	0.70	0.57	0.45	0.35

Note: A, B, and C denote the detrending procedure corresponding to the given model under the alternative hypothesis.

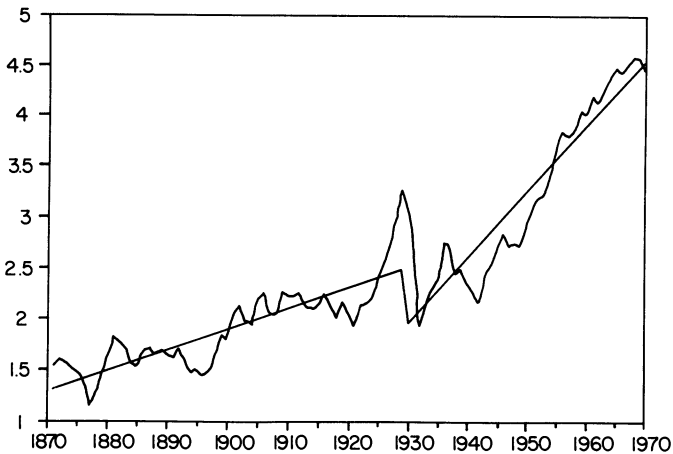


Note: The broken straight line is a fitted trend (by OLS) of the form:  $\tilde{y}_t = \tilde{\mu} + \tilde{\beta}t + \tilde{\gamma}DT_t^*$  where  $DT_t^* = 0$  if  $t \leq 1973:I$  and  $DT_t^* = t - T_B$  if  $t > 1973:I = T_B$ .

FIGURE 2.—Logarithm of "Postwar Quarterly Real GNP."

the full sample (given the quarterly nature of the series, the difference is important). The same features discussed above appear to hold when there is a change in the slope of the trend function.

As a final example, consider the common stock price series graphed in Figure 3. The break point is again in 1929 but in this case there appears to be both a sudden change in the level of the series in 1929 and a higher growth rate after. The solid line is the estimated trend with two dummy variables added, an intercept dummy (0 prior and at 1929, 1 after 1929) and a slope dummy (0 prior



Note: The broken straight line is a fitted trend (by OLS) of the form  $\tilde{y}_t = \tilde{\mu} + \tilde{\gamma}_1 DU_t + \tilde{\beta}t + \tilde{\gamma}_2 DT_t$  where  $DU_t = DT_t = 0$  if  $t \leq 1929$  and  $DU_t = 1$ ,  $DT_t = t$  if  $t > 1929$ .

FIGURE 3.—Logarithm of "Common Stock Prices."

and at 1929 and  $t$  after 1929). The estimated values of  $\alpha$  (in regression (1)) with the full sample are 0.913 but are only 0.732 using the pre-1929 sample and 0.788 using the post-1929. Here again, the  $t$  statistics are not large enough, however, to reject the unit root hypothesis at even the 10 percent level using any of the subsamples.<sup>2</sup>

Table II presents the autocorrelation function of the “detrended series” for the full set of variables analyzed by Nelson and Plosser, along with the postwar quarterly real GNP series. All series are detrended according to Model (A) (with a constant, a trend, and an intercept dummy) except for the postwar Quarterly Real GNP Series (with a slope dummy instead of the intercept dummy, Model (B)) and the real wage and common stock price series (with both a slope and intercept dummy, Model (C)). Unlike the “standard” detrended series (see Table 4 of Nelson-Plosser), the autocorrelations decay quite rapidly for all variables except for the consumer prices and velocity series. This behavior of the autocorrelation function is certainly not the one usually associated with either a random walk or a detrended random walk. Indeed, the “detrended” series appear stationary.

The results of this section motivate the analysis presented in the following sections. We first investigate the effects of the two types of changes in the trend function that we consider on the statistical properties of autoregressive estimates of the type found in regression (1) (both in finite samples and asymptotically). We find that such changes create a spurious unit root that may not vanish, even asymptotically. To overcome the problem of the low power associated with testing for a unit root using split samples, formal test statistics, which permit the presence of either or both an intercept and a slope shift, are developed in Section 4.

### 3. THE EFFECT OF A SHIFT IN THE TREND FUNCTION ON TESTS FOR A UNIT ROOT

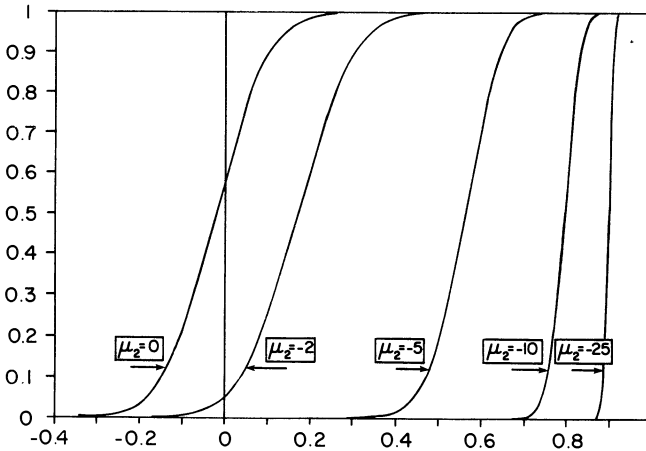
To assess the effects of the presence of a shift in the level of the series or a shift in the slope (at a single point of time) on tests for the presence of a unit root, we first present a small Monte Carlo experiment. Consider first the “crash hypothesis” (Model (A)). We generated 10,000 replications of a series  $\{y_t\}$  of length 100 defined by

$$(2) \quad y_t = \mu_1 + (\mu_2 - \mu_1)DU_t + \beta t + e_t \quad (t = 1, \dots, T),$$

where  $DU_t = 1$  if  $t > T_B$ , and 0 otherwise.

<sup>2</sup>Dickey-Fuller tests for the presence of a unit root using split samples are presented in Appendix B for all the series considered. The results are presented for values of  $k$  ranging from 1 to 12. These results show that (i) the conclusions drawn are not sensitive to the value of  $k$  chosen; (ii) for some series it is possible to reject the unit root hypothesis, especially when considering the post-1929 subsample. Furthermore, the statistical significance of the lagged first-differences (not reported) suggest that a large value of  $k$  may be needed. For example, the  $t$  statistics on the eighth lagged first-difference is often statistically significant. A similar pattern will occur in the full sample tests reported in subsequent sections.





Note:  $\tilde{\alpha}$  is the estimated autoregressive parameter in regression (4). The data-generating mechanism is given by equation (2) with  $\mu_1 = 0$ ,  $\beta = 1.0$  and  $\{e_t\}$  i.i.d.  $N(0, 1)$ ,  $T = 100$  and  $T_B = 50$ .

FIGURE 4.—C.D.F. of  $\tilde{\alpha}$  under the “Crash” Model.

For simplicity,  $\mu_1 = 0$ ,  $\beta = 1$ ,  $T_B = 50$ ,  $T = 100$  and the innovations  $e_t$  are i.i.d.  $N(0, 1)$ . For the “changing growth” hypothesis, a similar setup is considered except that  $y_t$  is generated by

$$(3) \quad y_t = \mu + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + e_t \quad (t = 1, \dots, T),$$

where  $DT_t^* = t - T_B$  if  $t > T_B$ , and 0 otherwise.

Again,  $\mu = 0$ ,  $\beta_1 = 1$ ,  $T_B = 50$ ,  $T = 100$ , and  $e_t \sim$  i.i.d.  $N(0, 1)$ . For each replication, we computed the autoregressive coefficient  $\tilde{\alpha}$  in the following regression, using ordinary least squares:

$$(4) \quad y_t = \tilde{\mu} + \tilde{\beta}t + \tilde{\alpha}y_{t-1} + \tilde{e}_t.$$

Figure 4 graphs the cumulative distribution function of  $\tilde{\alpha}$  when the data generating process (D.G.P.) is given by (2) for various values of  $\mu_2$ . This experiment reveals that as the magnitude of the crash increases ( $\mu_2$  decreases), the c.d.f. of  $\tilde{\alpha}$  becomes more concentrated at a value ever closer to 1. The corresponding mean and variance of the sample of  $\tilde{\alpha}$  generated are shown in Table III. Figure 5 graphs the c.d.f. of  $\tilde{\alpha}$  when the D.G.P. is given by (3) for various values of  $\beta_2$ . As  $\beta_2$  diverges from  $\beta_1$ , again, the c.d.f. becomes more concentrated and closer to one. The computed mean and variance of  $\tilde{\alpha}$  presented in Table III confirms this behavior.<sup>3</sup>

<sup>3</sup>Note that when the error structure is i.i.d.,  $\tilde{\alpha}$  is free of nuisance parameters and hence can be used as a formal test statistic on the same ground as the  $t$  statistic. However, we also performed a similar experiment with the  $t$  statistic on  $\tilde{\alpha}$  ( $\alpha = 1$ ) in regression (4) as well as in a regression with additional lags of first-differences as regressors. The results obtained show the same behavior. If anything, the  $t$  statistic with extra lags of first-differences as regressors shows a still greater bias toward nonrejection of the null hypothesis of a unit root. These results are available upon request. We prefer to report our result in terms of the behavior of the estimator  $\tilde{\alpha}$  instead of its  $t$  statistic because it makes clear that what causes the nonrejection is not due solely to the behavior of the variance estimator. What is of importance is that  $\tilde{\alpha}$  is biased towards unity.

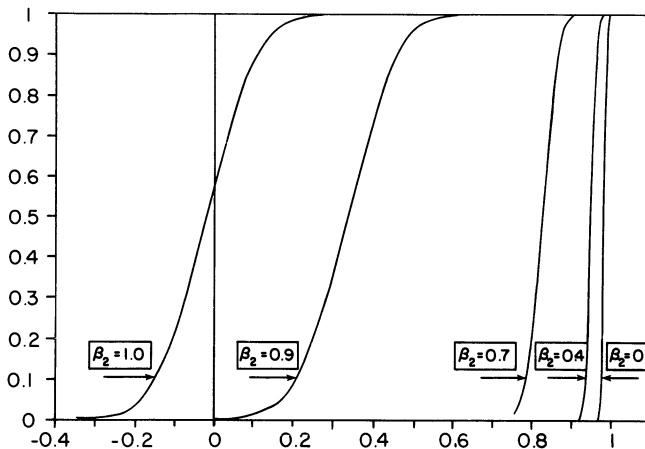
TABLE III  
MEAN AND VARIANCE OF  $\tilde{\alpha}$

(a) Crash Simulations, $\mu_1 = 0, \beta = 1$					
	$\mu_2 = 0$	$\mu_2 = -2$	$\mu_2 = -5$	$\mu_2 = -10$	$\mu_2 = -25$
Mean	-0.019	0.172	0.558	0.795	0.899
Variance	0.00986	0.01090	0.00471	0.00089	0.00009
(b) Breaking Trend Simulations, $\beta_1 = 1, \mu = 0$					
	$\beta_2 = 1.0$	$\beta_2 = 0.9$	$\beta_2 = 0.7$	$\beta_2 = 0.4$	$\beta_2 = 0.0$
Mean	-0.019	0.334	0.825	0.949	0.981
Variance	0.00986	0.00938	0.00094	0.00009	0.00001

See notes to Figure 4 for case (a) and Figure 5 for case (b).

What emerges from this experiment is that if the magnitude of the shift is significant, one could hardly reject the unit root hypothesis even if the series is that of a trend (albeit with a break) with i.i.d. disturbances. In particular, one would conclude that the shocks have permanent effects. Here, the shocks clearly have no permanent effects, only the one-time shift in the trend function is permanent.

To analyze the effect of an increase in the sample size on the distribution of  $\tilde{\alpha}$  with a shift of a given magnitude, we derive the asymptotic limit of  $\tilde{\alpha}$ . To this end, we again consider processes generated by Models (A), (B), or (C) under the alternative hypotheses, but we enlarge the framework by allowing general conditions on the error structure  $\{e_t\}$ . Many such sets of conditions are possible and would allow us to carry out the asymptotic theory. For simplicity, we use the



Note:  $\tilde{\alpha}$  is the estimated autoregressive parameter in regression (4). The data-generating mechanism is given by equation (3) with  $\mu = 0, \beta_1 = 1.0, \{e_t\}$  i.i.d.  $N(0, 1), T = 100, T_B = 50$ .

FIGURE 5.—C.D.F. of  $\tilde{\alpha}$  under the “Breaking Trend” Model.

“mixing-type” conditions of Phillips (1987) and Phillips and Perron (1988). These are stated in Assumption 1.

ASSUMPTION 1: (a)  $E(e_t) = 0$  all  $t$ ; (b)  $\sup_t E|e_t|^{\beta+\varepsilon} < \infty$  for some  $\beta > 2$  and  $\varepsilon > 0$ ; (c)  $\sigma^2 = \lim_{T \rightarrow \infty} T^{-1}E(S_T^2)$  exists and  $\sigma^2 > 0$ , where  $S_T = \sum_1^T e_i$ ; (d)  $\{e_t\}_1^\infty$  is strong mixing with mixing numbers  $\alpha_m$  that satisfy:  $\sum_1^\infty \alpha_m^{1-2/\beta} < \infty$ .

These conditions are general enough to permit the series  $\{e_t\}$  to be generated by a finite order ARMA( $p, q$ ) process with Gaussian innovations. To carry out the asymptotic analysis, we shall require that both the pre-break and post-break samples increase at the same rate as the total number of observations,  $T$ , increases. To this effect, we assume, for simplicity, that  $T_B = \lambda T$  for all  $T$ . We refer to  $\lambda$  as the “break fraction.” The asymptotic limits are taken as  $T$  increases to infinity in a sequence that ensures an integer value of  $T_B$  for a given  $\lambda$ . This type of increasing sequence is assumed throughout the paper. The results proved in Appendix A are presented in the following theorem.

THEOREM 1: Let  $\{y_t\}_0^T$  be a sample of size  $T + 1$  generated under one of the alternative hypotheses with the innovations  $\{e_t\}$  satisfying Assumption 1. Let ‘ $\rightarrow$ ’ denote convergence in probability. Furthermore, let  $T_B = \lambda T$ , for all  $T$  and  $0 < \lambda < 1$ ; then as  $T \rightarrow \infty$ :

(a) The “crash hypothesis”: Under Model (A)

$$\tilde{\alpha} \rightarrow \{[\mu_1 - \mu_2]^2 A + \gamma_1\} \{[\mu_1 - \mu_2]^2 A + \sigma_e^2\}^{-1}$$

where

$$A = [\lambda - 4\lambda^2 + 6\lambda^3 - 3\lambda^4], \quad \gamma_1 = \lim_{T \rightarrow \infty} T^{-1} \sum_1^T E(e_t e_{t-1}),$$

and

$$\sigma_e^2 = \lim_{T \rightarrow \infty} T^{-1} \sum_1^T E(e_t^2).$$

(b) The “breaking trend hypothesis”: Under either Model (B) or (C)

(i)  $\tilde{\alpha} \rightarrow 1,$

(ii)  $T(\tilde{\alpha} - 1) \rightarrow \{3(-1 + 4\lambda - 5\lambda^2 + 2\lambda^3)\} \cdot \{2(-3 + 4\lambda - 3\lambda^2 + 3\lambda^3 - \lambda^4)\}^{-1}.$

Part (a) of Theorem 1 shows that under the crash hypothesis, the limit of  $\tilde{\alpha}$  depends on the relative magnitude of  $[\mu_1 - \mu_2]^2 A$  and  $\sigma_e^2$ . In particular, this limit gets closer to one as  $[\mu_1 - \mu_2]^2$  increases. Another feature is that the limit of  $\tilde{\alpha}$  is always greater than the true first-order autoregressive coefficient of the stationary part of the series,  $\gamma_1/\sigma_e^2$ . However, since  $\tilde{\alpha}$  does not converge to 1, the

usual statistics for testing that  $\alpha = 1$ , such as  $T(\tilde{\alpha} - 1)$  or the  $t$  statistic on  $\tilde{\alpha}$ , would eventually reject the null hypothesis of a unit root. Nevertheless, added to the generally poor power properties of tests for a unit root is the consideration that the limit of  $\tilde{\alpha}$  is inflated above its true value. These conditions are such that it could be difficult to reject the unit root hypothesis in finite samples.

There is another interpretation to the results under the crash hypothesis. As stated in model (A), the change in the intercept of the trend function is given by  $(\mu_2 - \mu_1)$ , a fixed value. This implies that in the asymptotic derivations we are considering a shift which decreases relative to the level of the series as the sample size increases. It may be more appropriate to specify the change in the intercept as a magnitude relative to the level of the series at the time of the break. Since at this period the level of the series is proportional to  $T_B$ , we can specify  $(\mu_2 - \mu_1)$  as a proportion of  $T_B$ , say,  $(\mu_2 - \mu_1) = \gamma T_B$ . In this case the "crash" is proportional to the level of the series. Since  $T_B \rightarrow \infty$  as  $T \rightarrow \infty$ , it is clear, from part (a), that under this interpretation,  $\tilde{\alpha} \rightarrow 1$ .

Such ambiguity does not occur under the "breaking trend hypothesis" (Models (B) or (C)) as is shown by part (b) of Theorem 1. Here, the limit of  $\tilde{\alpha}$  is 1 irrespective of the behavior of the intercept and the limit of  $T(\tilde{\alpha} - 1)$  is invariant to the relative magnitude of the shift ( $\beta_2$  versus  $\beta_1$ ). The expression in part (b, ii) is a function of  $\lambda$ . However, it varies from 0 to 1/2 for values of  $\lambda$  in the range (0, 1). Since the one-sided 5 percent asymptotic critical value of  $T(\tilde{\alpha} - 1)$  is  $-21.8$  under the null hypothesis of a unit root (Fuller (1976)), Theorem 2 implies that the unit root hypothesis could not be rejected, even asymptotically.<sup>4</sup>

These results could be extended to more general test statistics, such as the  $t$  statistics. Nevertheless, the picture is clear. Tests of the unit root hypothesis are not consistent against "trend stationary" alternatives when the trend function contains a shift in the slope. Although they are not inconsistent against a shift in the intercept of the trend function (if the change is fixed as  $T$  increases), their power is likely to be substantially reduced due to the fact that the limit of the autoregressive coefficient is inflated above its true value. When interpreting the "crash" as proportional to the level of the series as  $T$  increases,  $\tilde{\alpha}$  unambiguously converges to one and implies a considerable loss in power. There is therefore a need to develop alternative statistical procedures that could distinguish a process with a unit root from a process stationary around a breaking trend function.

#### 4. ALTERNATIVE STATISTICAL PROCEDURES

In this section, we extend the Dickey-Fuller testing strategy to ensure a consistent testing procedure against shifting trend functions. We shall present several ways to do so, all of which are asymptotically equivalent, and discuss the main differences between each.

<sup>4</sup>After the first draft of this paper was written, we became aware of a result similar to part (b, i) of Theorem 1 proved by Rappoport and Reichlin (1987). In fact, in the case of deterministic trends with multiple shifts in slope, they prove the following more general result: "If the true model contains  $K + 1$  segments, then any fitted model involving  $K$  or less segments will, asymptotically, yield a larger sum of squared residuals than [a difference stationary] model" (p.9).

Consider first detrending the raw series  $\{y_t\}$  according to either model (A), (B), or (C). Let  $\{\tilde{y}_t^i\}$ ,  $i = A, B, C$  be the residuals from a regression of  $y_t$  on (1)  $i = A$ : a constant, a time trend, and  $DU_t$ ; (2)  $i = B$ : a constant, a time trend, and  $DT_t^*$ ; (3)  $i = C$ : a constant, a time trend,  $DU_t$ , and  $DT_t$ . Furthermore, let  $\tilde{\alpha}^i$  be the least squares estimator of  $\alpha$  in the following regression:

$$(5) \quad \tilde{y}_t^i = \tilde{\alpha}^i \tilde{y}_{t-1}^i + \tilde{e}_t \quad (i = A, B, C; t = 1, 2, \dots, T).$$

Up to this point the extensions from the no break model are straightforward enough. However, matters are not so simple concerning the distribution of the statistics of interest, namely the normalized bias,  $T(\tilde{\alpha}^i - 1)$ , and the  $t$  statistic on  $\tilde{\alpha}^i$ ,  $t_{\tilde{\alpha}^i}$  ( $i = A, B, C$ ). Needless to say, the only manageable analytical distribution theory is asymptotic in nature. But two additional features are also present over the usual procedure: (a) extra regressors and (b) the split sample nature of these extra regressors. To this effect, we derive the asymptotic distribution of  $T(\tilde{\alpha}^i - 1)$  and  $t_{\tilde{\alpha}^i}$  under the null hypothesis of a unit root. As in Section 3, we require that the break point  $T_B$  increases at the same rate as the total sample size  $T$ . Again, for simplicity, it is assumed that  $T_B = \lambda T$  with both  $T$  and  $T_B$  integer valued.

The method of proof is similar to that of Phillips (1987) and Phillips and Perron (1988). We use weak convergence results that hold for normalized functions of the sum of the innovations when the latter are assumed to satisfy Assumption 1. The limiting distributions obtained under this general setting are then specialized to the i.i.d. case. The asymptotic distributions in the i.i.d. case are evaluated using simulations, and critical values are tabulated. We then show how the results can be extended to innovations  $\{e_t\}$  that follow the general ARMA( $p, q$ ) process in the same way that the Dickey-Fuller regressions are modified by adding extra lags of first-differences of the data as regressors.

The main results concerning the asymptotic distributions of the normalized bias estimators and the  $t$  statistics of the autoregressive coefficient under the null hypothesis of a unit root are presented in the next theorem.

**THEOREM 2:** *Let  $\{y_t\}$  be generated under the null hypothesis of model  $i$  ( $i = A, B, C$ ) with the innovation sequence  $\{e_t\}$  satisfying Assumption 1. Let  $\Rightarrow$  denote weak convergence in distribution and  $\lambda = T_B/T$  for all  $T$ . Then, as  $T \rightarrow \infty$ :*

- (a)  $T(\tilde{\alpha}^i - 1) \Rightarrow H_i/K_i,$
- (b)  $t_{\tilde{\alpha}^i} \Rightarrow (\sigma/\sigma_e)H_i/(g_iK_i)^{1/2},$

where

$$\begin{aligned} H_A &= g_A D_1 - D_5 \psi_1 - D_6 \psi_2; & K_A &= g_A D_2 - D_4 \psi_2 - D_3 \psi_1; \\ H_B &= g_B D_1 + D_5 \psi_3 + D_8 \psi_4; & K_B &= g_B D_2 + D_7 \psi_4 + D_3 \psi_3; \\ H_C &= g_C D_9 + D_{13} \psi_5 - D_{14} \psi_6; & K_C &= g_C D_{10} - D_{12} \psi_6 + D_{11} \psi_5; \end{aligned}$$

with

$$\psi_1 = 6D_4 + 12D_3; \quad \psi_2 = 6D_3 + (1 - \lambda)^{-1}\lambda^{-1}D_4;$$

$$\psi_3 = (1 + 2\lambda)(1 - \lambda)^{-1}D_7 - (1 + 3\lambda)D_3;$$

$$\psi_4 = (1 + 2\lambda)(1 - \lambda)^{-1}D_3 - (1 - \lambda)^{-3}D_7;$$

$$\psi_5 = D_{12} - D_{11}; \quad \psi_6 = \psi_5 + (1 - \lambda)^2 D_{12}/\lambda^3;$$

and

$$D_1 = \left(\frac{1}{2}\right)(w(1)^2 - \sigma_e^2/\sigma^2) - w(1) \int_0^1 w(r) dr;$$

$$D_2 = \int_0^1 w(r)^2 dr - \left[ \int_0^1 w(r) dr \right]^2;$$

$$D_3 = \int_0^1 rw(r) dr - \left(\frac{1}{2}\right) \int_0^1 w(r) dr; \quad D_4 = \int_0^\lambda w(r) dr - \lambda \int_0^1 w(r) dr;$$

$$D_5 = w(1)/2 - \int_0^1 w(r) dr; \quad D_6 = w(\lambda) - \lambda w(1);$$

$$D_7 = \int_\lambda^1 rw(r) dr - \lambda \int_\lambda^1 w(r) dr - ((1 - \lambda)^2/2) \int_0^1 w(r) dr;$$

$$D_8 = ((1 - \lambda^2)/2)w(1) - \int_\lambda^1 w(r) dr;$$

$$D_9 = \int_0^1 w(r)^2 dr - \lambda^{-1} \left( \int_0^\lambda w(r) dr \right)^2 - (1 - \lambda)^{-1} \left( \int_\lambda^1 w(r) dr \right)^2;$$

$$D_{10} = (w(1)^2 - \sigma_e^2/\sigma^2)/2 - \lambda^{-1}w(\lambda) \int_0^\lambda w(r) dr \\ - (w(1) - w(\lambda))(1 - \lambda)^{-1} \int_\lambda^1 w(r) dr;$$

$$D_{11} = \int_0^1 rw(r) dr - \left(\frac{1}{2}\right)(1 + \lambda) \int_0^1 w(r) dr + \left(\frac{1}{2}\right) \int_0^\lambda w(r) dr;$$

$$D_{12} = \int_0^\lambda rw(r) dr - (\lambda/2) \int_0^\lambda w(r) dr;$$

$$D_{13} = (1 - \lambda)w(1)/2 + w(\lambda)/2 - \int_0^1 w(r) dr;$$

$$D_{14} = \lambda w(\lambda)/2 - \int_0^\lambda w(r) dr;$$

$$g_A = 1 - 3(1 - \lambda)\lambda; \quad g_B = 3\lambda^3; \quad g_C = 12(1 - \lambda)^2;$$

and where  $w(r)$  is the unit Wiener process defined on  $C[0, 1]$ ,  $\sigma^2 = \lim_{T \rightarrow \infty} E[T^{-1} S_T^2]$ ,  $S_T = \sum_1^T e_t$ , and  $\sigma_e^2 = \lim_{T \rightarrow \infty} E[T^{-1} \sum_1^T e_t^2]$ .

Theorem 2 provides a representation for the limiting distribution of the normalized least squares estimators and their  $t$  statistics in terms of functionals of Wiener processes. These limiting distributions are functions of the parameter  $\lambda$ , the ratio of the pre-break sample size to total sample size. It is easy to verify that when  $\lambda$  is either 0 or 1, the limiting distributions are identical over all models and are given by:

$$T(\tilde{\alpha} - 1) \Rightarrow H/K \quad \text{and} \quad t_{\tilde{\alpha}} \Rightarrow (\sigma/\sigma_e)H/K^{1/2}$$

where

$$\begin{aligned}
 H &= \left(\frac{1}{2}\right)(w(1)^2 - \sigma_e^2/\sigma^2) + 12 \left[ \int_0^1 r w(r) dr - \left(\frac{1}{2}\right) \int_0^1 w(r) dr \right] \\
 &\quad \cdot \left[ \int_0^1 w(r) dr - \left(\frac{1}{2}\right) w(1) \right] - w(1) \int_0^1 w(r) dr, \\
 K &= \int_0^1 w(r)^2 dr - 12 \left( \int_0^1 r w(r) dr \right)^2 \\
 &\quad + 12 \int_0^1 w(r) dr \int_0^1 r w(r) dr - 4 \left( \int_0^1 w(r) dr \right)^2.
 \end{aligned}$$

These latter asymptotic distributions correspond to those derived by Phillips and Perron (1988) in the case where no dummy variables are included.

The expressions for the limiting distributions in Theorem 2 depend on additional nuisance parameters, apart from  $\lambda$ , namely  $\sigma^2$  and  $\sigma_e^2$ . As in Phillips (1987) and Phillips and Perron (1988),  $\sigma_e^2$  is the variance of the innovations and  $\sigma^2$  is, in the case of weakly stationary innovations, equal to  $2\pi f(0)$  where  $f(0)$  is the spectral density of  $\{e_t\}$  evaluated at frequency zero. When the innovation sequence  $\{e_t\}$  is independent and identically distributed,  $\sigma^2 = \sigma_e^2$  and, in that case, the limiting distributions are invariant with respect to nuisance parameters, except  $\lambda$ .

Therefore, when  $\sigma^2 = \sigma_e^2$ , percentage points of the limiting distributions can be tabulated for given values of  $\lambda$ . Tables IV, V, and VI present selected percentage points that will allow us to carry hypothesis testing. The critical values are obtained via simulation methods. We briefly describe the steps involved. First, we generate a sample of size 1,000 of i.i.d.  $N(0, 1)$  random deviates,  $\{e_t\}$ . We then construct sample moments of the data which converge weakly to the various functionals of the Wiener process involved in the representation of the asymptotic distributions. For example, as  $T \rightarrow \infty$ ,  $T^{-1/2} \sum_1^T e_t \Rightarrow w(1)$ ,  $T^{-1/2} \sum_1^{T\lambda} e_t \Rightarrow w(\lambda)$ ,  $T^{-3/2} \sum_1^T \sum_{j=1}^t e_j \Rightarrow \int_0^1 w(r) dr$ ,  $\sum_{t=1}^T (\sum_{j=1}^{t-1} e_j) e_t \Rightarrow \left(\frac{1}{2}\right)(w(1)^2 - 1)$ , etc. With a sample size of 1,000 and i.i.d.  $N(0, 1)$  variates, we can expect the approximation to be quite accurate. Once the various functionals are evaluated, we construct the expressions in Theorem 2 and obtain one realization of the limiting distributions

TABLE IV.A  
PERCENTAGE POINTS OF THE ASYMPTOTIC DISTRIBUTION OF  $T(\tilde{\alpha} - 1)$  IN MODEL A  
Time of Break Relative to Total Sample Size:  $\lambda$

$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1%	-34.17	-35.85	-35.07	-34.44	-34.07	-35.83	-35.59	-34.86	-34.65
2.5%	-28.93	-30.35	-29.92	-29.26	-29.00	-29.80	-29.61	-29.40	-29.35
5%	-25.04	-26.00	-25.90	-25.40	-25.25	-25.56	-25.99	-25.82	-25.40
10%	-21.45	-22.16	-21.93	-21.61	-21.55	-21.79	-22.33	-22.10	-21.48
90%	-4.57	-5.19	-5.13	-4.28	-3.85	-4.36	-5.15	-5.32	-4.62
95%	-3.40	-3.90	-3.80	-2.83	-2.38	-2.92	-3.86	-3.87	-3.27
97.5%	-2.35	-2.92	-2.85	-1.69	-1.42	-1.89	-2.78	-2.84	-2.13
99%	-1.28	-1.70	-1.60	-0.61	-0.40	-0.78	-1.58	-1.78	-1.39

TABLE IV.B  
PERCENTAGE POINTS OF THE ASYMPTOTIC DISTRIBUTION OF  $t_{\tilde{\alpha}}$  IN MODEL A  
Time of Break Relative to Total Sample Size:  $\lambda$

$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1%	-4.30	-4.39	-4.39	-4.34	-4.32	-4.45	-4.42	-4.33	-4.27
2.5%	-3.93	-4.08	-4.03	-4.01	-4.01	-4.09	-4.07	-3.99	-3.97
5%	-3.68	-3.77	-3.76	-3.72	-3.76	-3.76	-3.80	-3.75	-3.69
10%	-3.40	-3.47	-3.46	-3.44	-3.46	-3.47	-3.51	-3.46	-3.38
90%	-1.38	-1.45	-1.43	-1.26	-1.17	-1.28	-1.42	-1.46	-1.37
95%	-1.09	-1.14	-1.13	-0.88	-0.79	-0.92	-1.10	-1.13	-1.04
97.5%	-0.78	-0.90	-0.83	-0.55	-0.49	-0.60	-0.82	-0.89	-0.74
99%	-0.46	-0.54	-0.51	-0.21	-0.15	-0.26	-0.50	-0.57	-0.47

TABLE V.A  
PERCENTAGE POINTS OF THE ASYMPTOTIC DISTRIBUTION OF  $T(\tilde{\alpha} - 1)$  IN MODEL B  
Time of Break Relative to Total Sample Size:  $\lambda$

$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1%	-34.34	-37.16	-38.07	-39.21	-39.77	-40.08	-38.70	-36.18	-34.69
2.5%	-28.74	-31.97	-32.78	-33.42	-33.60	-33.21	-32.31	-31.45	-29.42
5%	-25.00	-27.16	-28.61	-29.23	-29.65	-29.51	-28.68	-27.24	-25.25
10%	-21.26	-23.10	-24.20	-25.04	-25.40	-25.15	-24.30	-23.01	-21.24
90%	-4.27	-5.09	-5.92	-6.62	-6.96	-6.71	-6.08	-5.26	-4.45
95%	-3.12	-3.85	-4.50	-5.06	-5.31	-5.15	-4.59	-3.82	-3.16
97.5%	-2.20	-2.69	-3.30	-3.90	-4.14	-3.94	-3.36	-2.72	-2.21
99%	-1.11	-1.58	-2.19	-2.50	-3.01	-2.54	-2.20	-1.50	-1.24

of the statistics  $T(\tilde{\alpha}^i - 1)$ ,  $t_{\tilde{\alpha}^i}$  ( $i = A, B, C$ ). We replicate this procedure 5,000 times and obtain the critical values from the sorted vector of replicated statistics. This procedure is performed for each statistic with nine values of the parameter  $\lambda$ , the ratio of pre-break sample size to total sample size.<sup>5</sup>

Several features are worth mentioning with respect to these critical values. First, as expected, for a given size of the test, the critical values are larger (in

<sup>5</sup>For some evidence on the adequacy of this method to obtain critical values for limiting distributions involving functions of Wiener processes, see Chan (1988).



TABLE V.B  
 PERCENTAGE POINTS OF THE ASYMPTOTIC DISTRIBUTION OF  $t_{\hat{\alpha}}$  IN MODEL B  
 Time of Break Relative to Total Sample Size:  $\lambda$

$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1%	-4.27	-4.41	-4.51	-4.55	-4.56	-4.57	-4.51	-4.38	-4.26
2.5%	-3.94	-4.08	-4.17	-4.20	-4.26	-4.20	-4.13	-4.07	-3.96
5%	-3.65	-3.80	-3.87	-3.94	-3.96	-3.95	-3.85	-3.82	-3.68
10%	-3.36	-3.49	-3.58	-3.66	-3.68	-3.66	-3.57	-3.50	-3.35
90%	-1.35	-1.48	-1.59	-1.69	-1.74	-1.71	-1.61	-1.49	-1.34
95%	-1.04	-1.18	-1.27	-1.37	-1.40	-1.36	-1.28	-1.16	-1.04
97.5%	-0.78	-0.87	-0.97	-1.11	-1.18	-1.11	-0.97	-0.87	-0.77
99%	-0.40	-0.52	-0.69	-0.75	-0.82	-0.78	-0.67	-0.54	-0.43

TABLE VI.A  
 PERCENTAGE POINTS OF THE ASYMPTOTIC DISTRIBUTION OF  $T(\hat{\alpha} - 1)$  IN MODEL C  
 Time of Break Relative to Total Sample Size:  $\lambda$

$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1%	-36.17	-39.97	-42.98	-45.52	-44.07	-44.75	-43.02	-41.48	-36.58
2.5%	-30.65	-34.92	-36.48	-37.12	-37.56	-37.72	-37.50	-35.16	-31.82
5%	-26.63	-29.95	-32.47	-33.22	-33.79	-33.19	-33.11	-30.70	-27.16
10%	-22.68	-25.50	-27.90	-29.39	-29.41	-29.04	-28.14	-25.79	-22.62
90%	-4.74	-5.85	-7.35	-8.43	-8.84	-8.55	-7.41	-6.17	-4.89
95%	-3.41	-4.34	-5.50	-6.67	-7.19	-6.79	-5.66	-4.52	-3.52
97.5%	-2.51	-3.19	-4.14	-5.37	-5.82	-5.47	-4.33	-3.35	-2.49
99%	-1.31	-2.14	-2.82	-3.96	-4.39	-4.24	-2.80	-2.02	-1.28

TABLE VI.B  
 PERCENTAGE POINTS OF THE ASYMPTOTIC DISTRIBUTION OF  $t_{\hat{\alpha}}$  IN MODEL C  
 Time of Break Relative to Total Sample Size:  $\lambda$

$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1%	-4.38	-4.65	-4.78	-4.81	-4.90	-4.88	-4.75	-4.70	-4.41
2.5%	-4.01	-4.32	-4.46	-4.48	-4.53	-4.49	-4.44	-4.31	-4.10
5%	-3.75	-3.99	-4.17	-4.22	-4.24	-4.24	-4.18	-4.04	-3.80
10%	-3.45	-3.66	-3.87	-3.95	-3.96	-3.95	-3.86	-3.69	-3.46
90%	-1.44	-1.60	-1.78	-1.91	-1.96	-1.93	-1.81	-1.63	-1.44
95%	-1.11	-1.27	-1.46	-1.62	-1.69	-1.63	-1.47	-1.29	-1.12
97.5%	-0.82	-0.98	-1.15	-1.35	-1.43	-1.37	-1.17	-1.04	-0.80
99%	-0.45	-0.67	-0.81	-1.04	-1.07	-1.08	-0.79	-0.64	-0.50

absolute value) for each model than the standard Dickey-Fuller critical values when considering the left tail. One would therefore expect a loss in power. Secondly, although the critical values are not significantly influenced by the value of the parameter  $\lambda$ , the maximum (in absolute value) occurs around the value  $\lambda = 0.5$ , i.e., for a break at mid-sample.<sup>6</sup> In the left tail, critical values of the statistics are smallest (in absolute value) when  $\lambda$  is close to 0 or 1. This is to be

<sup>6</sup>The simulated critical values suggest that the limiting distributions are symmetric around  $\lambda = 0.5$ . This feature seems intuitively plausible. We have not, however, proved that such is the case.

expected since, as previously mentioned, the critical values are identical to those of Dickey and Fuller when  $\lambda = 0, 1$ .

Some critical values are worth noticing. Consider the  $t$  statistics. Under Model (A), the “crash hypothesis”, the 5 percent critical value has a minimum (over values of  $\lambda$ ) of  $-3.80$ . Under Models (B) and (C), the corresponding figures are  $-3.96$  and  $-4.24$  respectively. The critical values under the various models are therefore noticeably smaller than the standard Dickey-Fuller critical value of  $-3.41$  (see Dickey (1976) and Fuller (1976)).

These sets of results can be used to perform hypothesis testing. One simply picks the critical value corresponding to the sample value of  $\lambda$  at the chosen significance level. Since we only provide critical values for a selected grid of  $\lambda$ 's, the procedure suggested is to choose the critical value corresponding to the value of  $\lambda$  nearest to its sample value, i.e.,  $T_B/T$ . Given that the differences in the critical values over adjacent values for  $\lambda$  in the tables are not substantially different, this procedure should not cause misleading inferences.

4.1. *Extensions to More General Error Processes: Case (1)*

Using regressions (5) ( $i = A, B, C$ ) and the critical values in Tables IV, V, and VI is valid only in the case where the innovation sequence  $\{e_t\}$  is uncorrelated. When there is additional correlation, as one would expect, an extension is necessary. Two approaches are possible. One is to follow the approach suggested by Phillips (1987) and Phillips and Perron (1988). This involves finding a set of transformed statistics that would converge weakly to the limiting distributions expressed in Theorem 2 with  $\sigma^2 = \sigma_e^2$ . The other approach is that suggested by Dickey and Fuller (1979) and Said and Dickey (1984).

Consider first the extension to Phillips' (1987) procedure. It is useful first to write the limiting distributions of Theorem 2 in a different, more compact form. To do so, we adopt the framework suggested by Ouliaris, Park, and Phillips (1988). Define  $w_i(r)$  ( $i = A, B, C$ ) to be the stochastic process on  $[0, 1]$  such that  $w_i(r)$  is the projection residual of a Wiener process  $w(r)$  on the subspace generated by the following set of functions: (1)  $i = A: 1, r, du(r)$ , where  $du(r) = 0$  if  $r \leq \lambda$  and  $du(r) = 1$  if  $r > \lambda$ ; (2)  $i = B: 1, r, dt^*(r)$ , where  $dt^*(r) = 0$  if  $r \leq \lambda$  and  $dt^*(r) = r - \lambda$  if  $r > \lambda$ ; (3)  $i = C: 1, r, du(r), dt(r)$ , where  $dt(r) = 0$  if  $r \leq \lambda$  and  $dt(r) = r$  if  $r > \lambda$ . Adopting this notation, an alternative representation of the limiting distributions in Theorem 2 is given by:

$$T(\tilde{\alpha}^i - 1) \Rightarrow \left( \int_0^1 w_i(r) dw(r) + \delta \right) \left( \int_0^1 w_i(r)^2 dr \right)^{-1} \quad (i = A, B, C)$$

and

$$t_{\tilde{\alpha}^i} \Rightarrow (\sigma/\sigma_e) \left( \int_0^1 w_i(r) dw(r) + \delta \right) \left( \int_0^1 w_i(r)^2 dr \right)^{-1/2} \quad (i = A, B, C)$$

where  $\delta = (\sigma^2 - \sigma_e^2)/(2\sigma^2)$ .

Now, define  $\tilde{\sigma}_i^2$  and  $\tilde{\sigma}_{ei}^2$  to be, respectively, any consistent estimator of  $\sigma^2$  and  $\sigma_e^2$  based on the estimated residuals from regression (5) ( $i = A, B, C$ ).<sup>7</sup> Also define  $S_i^2$  ( $i = A, B, C$ ) to be the residual sum of squares from the regression  $y_{t-1}$  on (1)  $i = A$ :  $1, t, DU_t$ ; (2)  $i = B$ :  $1, t, DT_t^*$ ; and (3)  $i = C$ :  $1, t, DU_t, DT_t$ . We then define the transformed statistics as:

$$(6) \quad Z(\tilde{\alpha}^i) = T(\tilde{\alpha}^i - 1) - T^2(\tilde{\sigma}_i^2 - \tilde{\sigma}_{ei}^2)/2S_i^2 \quad (i = A, B, C),$$

$$(7) \quad Z(t_{\tilde{\alpha}^i}) = (\tilde{\sigma}_{ei}/\tilde{\sigma}_i)t_{\tilde{\alpha}^i} - T(\tilde{\sigma}_i^2 - \tilde{\sigma}_{ei}^2)/2\tilde{\sigma}_i S_i \quad (i = A, B, C).$$

Following Ouliaris, Park, and Phillips (1988), it is straightforward to show that:

$$(8) \quad Z(\tilde{\alpha}^i) \Rightarrow \left( \int_0^1 w_i(r) dw(r) \right) \left( \int_0^1 w_i(r)^2 dr \right)^{-1} \quad (i = A, B, C)$$

and

$$(9) \quad Z(t_{\tilde{\alpha}^i}) \Rightarrow \left( \int_0^1 w_i(r) dw(r) \right) \left( \int_0^1 w_i(r)^2 dr \right)^{-1/2} \quad (i = A, B, C).$$

The limiting distributions in (8) and (9) are those whose critical values are presented in Tables IV, V, and VI derived using the representation given by Theorem 2.

The other approach adopts the procedures suggested by Dickey and Fuller (1979, 1981) and Said and Dickey (1984) which add extra lags of the first differences of the data as regressors in equation (5). This extended framework is characterized by the following regression (again estimated by OLS):

$$(10) \quad \tilde{y}_t^i = \tilde{\alpha}^i \tilde{y}_{t-1}^i + \sum_{j=1}^k \tilde{c}_j \Delta \tilde{y}_{t-j}^i + \tilde{e}_t \quad (i = A, B, C)$$

where

$$\Delta \tilde{y}_t^i = \tilde{y}_t^i - \tilde{y}_{t-1}^i.$$

In the above representation,  $\tilde{\alpha}$  is the OLS estimator of  $\alpha$ , the sum of the autoregressive coefficients and the test is again that  $\alpha = 1$ . The parameter  $k$  specifies the number of extra regressors added. In a simple AR( $p$ ) process,  $k = p$ . In a more general ARMA( $p, q$ ) process with  $p$  and  $q$  unknown,  $k$  must increase at a controlled rate with the sample size. Arguments similar to those developed by Said and Dickey can be used to show that the limiting distributions of the statistics  $t_{\tilde{\alpha}^i}$  ( $i = A, B, C$ ) are the same when the innovation sequence is an ARMA( $p, q$ ) process and regression (10) is used, as they are when the errors are i.i.d. and regression (5) is used. However, slightly more restrictive assumptions are needed with respect to the innovation sequence  $\{e_t\}$  and the truncation parameter  $k$  for this asymptotic equivalence to hold. They are detailed in the following Assumption (see Said and Dickey (1984)).

<sup>7</sup>See Phillips (1987), Phillips and Perron (1988), and Perron (1988) for details.

ASSUMPTION 2: (a)  $A(L)e_t = B(L)v_t$ ; (b)  $v_t$  is a sequence of i.i.d.  $(0, \sigma^2)$  random variables with finite  $(4 + \delta)$ th moment for some  $\delta > 0$ ; (c)  $k \rightarrow \infty$  and  $T^{-1}k^3 \rightarrow 0$  as  $T \rightarrow \infty$ .

4.2. Extensions to More General Error Processes: Case (2)

A possible drawback of the methods suggested above is that they imply that the change in the trend function occurs instantaneously. Given, for instance, that the Great Depression was not an instantaneous event but lasted several years, one may wish to allow for such a transition period during changes in the trend function. One way to model this is to suppose that the economy reacts gradually to a shock to the trend function.<sup>8</sup> Consider, for instance, Model (A) where a crash occurs. A plausible specification of the trend function, say  $\eta_t^A$ , is given by:

$$(11) \quad \eta_t^A = \mu_1 + \beta t + \gamma \psi(L) DU_t$$

where  $\psi(L)$  is a stationary and invertible polynomial in  $L$  with  $\psi(0) = 1$  and  $\gamma = \mu_2 - \mu_1$ . The long run change in the trend function is given by  $\gamma\psi(1)$  while the immediate impact is simply  $\gamma$ . A similar framework holds for models (B) and (C).

One way to incorporate such a gradual change in the trend function is to suppose that the economy responds to a shock to the trend function the same way as it reacts to any other shock, i.e. to impose  $\psi(L) = B(L)^{-1}A(L)$  (see Section 2). In the literature on outliers specification the framework suggested here is analogous to the so-called ‘‘innovational outlier’’ model whereas the framework considered in Section 4.1 is analogous to the ‘‘additive outlier’’ model (see, e.g., Tsay (1986)). We can then implement tests for the presence of a unit root in a framework that directly extends the Dickey-Fuller strategy by adding dummy variables in regression (1). The following regressions, corresponding to Models (A), (B), and (C) are constructed by nesting the corresponding models under the null and alternative hypotheses:

$$(12) \quad y_t = \hat{\mu}^A + \hat{\theta}^A DU_t + \hat{\beta}^A t + \hat{d}^A D(TB)_t + \hat{\alpha}^A y_{t-1} + \sum_{i=1}^k \hat{c}_i \Delta y_{t-i} + \hat{e}_t,$$

$$(13) \quad y_t = \hat{\mu}^B + \hat{\theta}^B DU_t + \hat{\beta}^B t + \hat{\gamma}^B DT_t^* + \hat{\alpha}^B y_{t-1} + \sum_{i=1}^k \hat{c}_i \Delta y_{t-i} + \hat{e}_t,$$

$$(14) \quad y_t = \hat{\mu}^C + \hat{\theta}^C DU_t + \hat{\beta}^C t + \hat{\gamma}^C DT_t + \hat{d}^C D(TB)_t + \hat{\alpha}^C y_{t-1} + \sum_{i=1}^k \hat{c}_i \Delta y_{t-i} + \hat{e}_t.$$

The null hypothesis of a unit root imposes the following restrictions on the true parameters of each model: Model (A), the ‘‘crash hypothesis’’:  $\alpha^A = 1, \beta^A = 0,$

<sup>8</sup>Again, this treatment is analogous to the methodology proposed by Box and Tiao (1975) concerning intervention analyses.

$\theta^A = 0$ ; Model (B), the “breaking slope with no crash”:  $\alpha^B = 1$ ,  $\gamma^B = 0$ ,  $\beta^B = 0$ ; and Model (C), where both effects are allowed:  $\alpha^C = 1$ ,  $\gamma^C = 0$ ,  $\beta^C = 0$ . Under the alternative hypothesis of a “trend stationary” process, we expect  $\alpha^A, \alpha^B, \alpha^C < 1$ ;  $\beta^A, \beta^B, \beta^C \neq 0$ ;  $\theta^A, \theta^C, \gamma^B, \gamma^C \neq 0$ . Finally, under the alternative hypothesis,  $d^A, d^C$ , and  $\theta^B$  should be close to zero while under the null hypothesis they are expected to be significantly different from zero.

The asymptotic distribution of the  $t$  statistics  $t_{\hat{\alpha}}A$  and  $t_{\hat{\alpha}}C$  in (12) and (14) are the same, respectively, as the asymptotic distribution of  $t_{\tilde{\alpha}}A$  and  $t_{\tilde{\alpha}}C$  in (10). However such a correspondence does not hold for the  $t$  statistic  $t_{\hat{\alpha}}B$  in (13). Apart from the one-time dummy variable  $d(TB)_t$ , regressions (13) and (14) are equivalent; hence the asymptotic distribution of  $t_{\hat{\alpha}}B$  is identical to the asymptotic distribution of  $t_{\tilde{\alpha}}C$ . This implies that in the above framework, it is not possible to test for a unit root under the maintained hypothesis that the trend function has a change in slope with the two segments joined at the time of the change. Consequently, tests for the presence of a unit root in Model (B) will have less power using regression (13) than using (10) where the asymptotic critical values of the  $t$  statistic on  $\alpha$  are smaller (in absolute value). However, it is still possible to test for a unit root with constant drift against a trend stationary process as in Model (B) and use the critical values of Table V. One simply runs the regression:

$$(15) \quad y_t = \hat{\mu}^B + \hat{\beta}^B t + \hat{\gamma}^B DT_t^* + \hat{\alpha}^B y_{t-1} + \sum_{i=1}^k \hat{c}_i \Delta y_{t-i} + \hat{\epsilon}_t.$$

The asymptotic distribution of the  $t$  statistic  $t_{\hat{\alpha}}B$  in (15) is the same as the asymptotic distribution of the  $t$  statistic  $t_{\tilde{\alpha}}B$  in (10). Given that the regressor  $DU_t$  is absent from (15), this case, however, implies that the change in drift is not permitted under the null hypothesis.

Finally, note that it is possible to apply Phillips' nonparametric procedure using regressions (12) through (14) without the lagged first-differenced regressors and applying the corrections given by (6) and (7). However, such a procedure has the unattractive feature of imposing only a one-period adjustment to the change in the trend function. In the notation of (11), it imposes  $\psi(L) = 1 - \psi_1 L$  where  $\psi_1$  is the coefficient on the first lag in the polynomial  $B(L)^{-1}A(L)$ .

The procedures outlined in this section permit testing for the presence of a unit root in a quite general time series process which allows a one-time break in the mean of the series or its rate of growth (or both). In the next section, we apply these procedures in the specific context of breaks at the time of the 1929 crash and the 1973 oil shock.

## 5. EMPIRICAL APPLICATIONS

We apply the test statistics derived in the previous section to the data set used by Nelson and Plosser and to the postwar quarterly real GNP series. The data set considered by Nelson and Plosser consists of fourteen major macroeconomic series sampled at an annual frequency. We omit the analysis of the unemploy-

ment rate series given that it is generally perceived as being stationary. The sample varies for each series with a starting date between 1869 and 1909. However, each series ends in 1970. Given that we entertain the hypothesis that only the 1929 Great Crash and the 1973 oil price shock caused a major change in trend function, each series in this data set contains only one break. It is therefore possible to apply the tests described in the previous section. Similarly, the quarterly postwar real GNP series contains a single break as the sample goes from 1947:I to 1986:III. Following Nelson and Plosser, we consider the logarithm of each series except for the interest rate for which we use the level.

Of the thirteen series in the Nelson-Plosser data set that we analyze, preliminary investigations showed that eleven were potentially well-characterized by a trend function with a constant slope but with a major change in their level occurring right after the year 1929. For these series, the maintained hypothesis is, therefore, that of Model (A) and given that the Great Crash did not occur instantaneously but lasted several years, we apply regression (12) to carry out our testing procedure. The two series that were not modeled as such are the "real wages" and "common stock price" series. For these series, it appeared that not only a change in the level occurred after 1929 but there was also an increase in the slope of the trend function after this date. For these reasons, the maintained hypothesis is that of Model (C), and we use regression (14) to implement our tests.

The postwar quarterly real GNP series offers yet a different picture. The 1973 oil price shock did not cause a significant drop in the level of the series. However, after that date, the slope of the trend function has sensibly decreased. This phenomenon is consistent with the much discussed slowdown in the growth rate of real GNP since the mid-seventies; see, for example, the recent symposium in the *Journal of Economic Perspectives* (1988). For these reasons, the maintained hypothesis is that of Model (B). Given the inherent difficulty in testing for a unit root allowing lagged effects for the change in the trend function, we apply regression (10) ( $i = B$ ) to carry the testing procedure. Modeling the change in the trend function following the 1973 oil price shock as instantaneous, at least appears more plausible than the change that occurred during the Great Depression.

Table VII presents the corresponding estimated regressions for each series along with the  $t$  statistic on the parameters for the following respective hypotheses:  $\mu = 0$ ,  $\beta = 0$ ,  $\theta = 0$ ,  $\gamma = 0$ ,  $d = 0$ , and  $\alpha = 1$ . Recall that under the hypothesis of a unit root process  $\mu \neq 0$  (in general),  $\beta = 0$ ,  $\theta = 0$  (except in regression Models (C)),  $\gamma = 0$ ,  $d \neq 0$ , and  $\alpha = 1$ . Under the alternative hypothesis of stationary fluctuations around a deterministic breaking trend function:  $\mu \neq 0$ ,  $\theta \neq 0$ ,  $\beta \neq 0$ ,  $\gamma \neq 0$  (in general),  $d = 0$ , and  $\alpha < 1$ .

The value of  $k$  chosen is determined by a test on the significance of the estimated coefficients  $\hat{\epsilon}_i$ . We actually used a fairly liberal procedure choosing a value of  $k$  equal to say  $k^*$  if the  $t$  statistic on  $\hat{\epsilon}_i$  was greater than 1.60 in absolute value and the  $t$  statistic on  $\hat{\epsilon}_i$  for  $l > k^*$  was less than 1.60 (with a maximum value for  $k$  of 8, except for the postwar quarterly real GNP series

TABLE VII  
TESTS FOR A UNIT ROOT

(a) Regression (12), Model A; $y_t = \mu + \theta D U_t + \hat{\beta} t + \hat{\alpha} y_{t-1} + \hat{\delta} D(TB)_t + \hat{\alpha} y_{t-1} + \sum_{i=1}^k \hat{c}_i \Delta y_{t-i} + \hat{\epsilon}_t$														
$T_B = 1929$	$T$	$\lambda$	$k$	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\theta}$	$t_{\hat{\theta}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\delta}$	$t_{\hat{\delta}}$	$S(\hat{\epsilon})$		
Real GNP	62	0.33	8	3.441	5.07	-0.189	-4.28	0.0267	5.05	-0.018	-0.30	0.282	-5.03 <sup>a</sup>	0.0509
Nominal GNP	62	0.33	8	5.692	5.44	-0.360	-4.77	0.0359	5.44	0.100	1.09	0.471	-5.42 <sup>a</sup>	0.0694
Real per capita GNP	62	0.33	7	3.325	4.11	-0.102	-2.76	0.0111	4.00	-0.070	-1.09	0.531	-4.09 <sup>b</sup>	0.0555
Industrial production	111	0.63	8	0.120	4.37	-0.298	-4.58	0.0323	5.42	-0.095	-0.99	0.322	-5.47 <sup>a</sup>	0.0875
Employment	81	0.49	7	3.402	4.54	-0.046	-2.65	0.0057	4.26	-0.025	-0.77	0.667	-4.51 <sup>a</sup>	0.0295
GNP deflator	82	0.49	5	0.669	4.09	-0.098	-3.16	0.0070	4.01	0.026	0.53	0.776	-4.04 <sup>b</sup>	0.0438
Consumer prices	111	0.63	2	0.065	1.12	-0.004	-0.21	0.0005	1.75	-0.036	-0.79	0.978	-1.28	0.0445
Wages	71	0.41	7	2.38	5.45	-0.190	-4.32	0.0197	5.37	0.085	1.36	0.619	-5.41 <sup>a</sup>	0.0532
Money stock	82	0.49	6	0.301	4.72	-0.071	-2.59	0.0121	4.18	0.033	0.68	0.812	-4.29 <sup>b</sup>	0.0440
Velocity	102	0.59	0	0.050	0.932	-0.005	-0.20	-0.0002	-0.35	-0.136	-2.01	0.941	-1.66	0.0663
Interest rate	71	0.41	2	-0.018	-0.088	-0.343	-2.06	0.0105	2.64	0.197	0.64	0.976	-0.45	0.2787

(b) Regression (14), Model C; $y_t = \hat{\mu} + \hat{\theta} D U_t + \hat{\beta} t + \hat{\gamma} D(TB)_t + \hat{\alpha} y_{t-1} + \sum_{i=1}^k \hat{c}_i \Delta y_{t-i} + \hat{\epsilon}_t$																
$T_B = 1929$	$T$	$\lambda$	$k$	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\theta}$	$t_{\hat{\theta}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\delta}$	$t_{\hat{\delta}}$	$S(\hat{\epsilon})$		
Common stock prices	100	0.59	1	0.353	4.09	-1.051	-4.29	0.0070	4.43	0.0139	3.98	0.128	0.76	0.718	-4.87 <sup>b</sup>	0.1402
Real wages	71	0.41	8	2.115	4.33	-0.190	-3.71	0.0107	3.79	0.0066	3.33	0.031	0.78	0.298	-4.28 <sup>c</sup>	0.0330

(c) Regression (10), Model B; $y_t = \hat{\mu} + \hat{\beta} t + \hat{\gamma} D(TB)_t + \hat{\alpha} y_{t-1} + \sum_{i=1}^k \hat{c}_i \Delta y_{t-i} + \hat{\epsilon}_t$												
$T_B = 1973:1$	$T$	$\lambda$	$k$	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$S(\hat{\epsilon})$
Quarterly real GNP	159	0.66	10	6.977	1160.51	0.0087	97.73	-0.0031	-12.06	0.86	-3.98 <sup>c</sup>	0.0097

NOTE: a, b, and c denote statistical significance at the 1%, 2.5%, and 5% level respectively.

where we used a maximum of 12). This liberal procedure is justified in the sense that including too many extra regressors of lagged first-differences does not affect the size of the test but only decreases its power. Including too few lags may have a substantial effect on the size of the test.

Consider first the series for which we applied Model (A). To evaluate the significance of the  $t$  statistic on  $\hat{\alpha}$ , we use the critical value presented in Table IV.B with a value of  $\lambda$  closest to the ratio of pre-break sample size to total sample size. Of the eleven series in that group, the unit root hypothesis cannot be rejected even at the 10 percent level for three of them: "consumer prices," "interest rate" and "velocity." However, we can reject the null hypothesis of a unit root at that 2.5 percent level or better for all other eight series. We can reject it at the 1 percent level for the following series: "real GNP," "nominal GNP," "industrial production," "employment," and "wages," and at the 2.5 percent level for the series "real per capita GNP," "GNP deflator," and "money stock." In some cases the coefficient  $\hat{\alpha}$ , which is an estimate of the sum of the autoregressive coefficients, is dramatically different from one. For example, it is 0.282 for "real GNP" and 0.322 for "industrial production."

Given that the unit root hypothesis can be rejected for the eight series mentioned above, we can assess the significance of the other coefficients using the fact that the asymptotic distribution of their  $t$  statistic is standardized normal. In all cases, the estimated coefficients on the constant ( $\hat{\mu}$ ), the post-break dummy ( $\hat{\theta}$ ), and the trend ( $\hat{\beta}$ ) are significant at least at the 5 percent level. All series showed a trend function with a positive slope and a significant decrease in level just after 1929. For these eight series the coefficient on the break dummy ( $\hat{d}$ ) is not significant. These results strongly suggest that, except for the "consumer price," "velocity," and "interest rate" series, the underlying process is one of stationary fluctuations around a deterministic trend function.

Consider now panel (b) of Table VII which presents the results for the "common stock price," and "real wages" series estimated under Model (C). We can reject the null hypothesis of a unit root at the 2.5 percent level for "common stock prices" and at the 5 percent level for the "real wages" series. In both cases, the constant ( $\hat{\mu}$ ), the post-break constant dummy ( $\hat{\theta}$ ), the trend ( $\hat{\beta}$ ), and the post-break slope dummy ( $\hat{\gamma}$ ) are highly significant, while the break dummy ( $\hat{d}$ ) is not. The coefficients  $\hat{\alpha}$  for the "real wages" series is very low at 0.298 while for the "common stock price" it is at 0.718 showing substantial mean reversion effects. This finding about the "common stock price" series is particularly striking given the vast amount of theoretical and empirical studies supporting the random walk hypothesis in this situation.

Finally, panel (c) of Table VII presents the results for the postwar quarterly real GNP series using regression (10) corresponding to Model (B). In this case, the null hypothesis that  $\alpha = 1$  can be rejected at the 5 percent level with an estimated coefficient  $\hat{\alpha}$  equal to 0.86. This result is especially significant given the usual poor power properties of tests for a unit root against stationary alternatives when using a data set with a small span sampled frequently (see, e.g., Perron (1987) and Shiller and Perron (1985)). The other estimated coefficients in panel (c) confirm the relevance of the "trend stationary" model versus the "unit root" model. The coefficient on the post-break slope dummy coefficient ( $\hat{\gamma}$ ) is highly



significant. The estimated regression is therefore consistent with an underlying process characterized by stationary fluctuations around a deterministic trend function with a decrease in the slope after 1973.<sup>9</sup>

Table A3 in Appendix B presents the estimated value for the sum of the autoregressive coefficient,  $\alpha$ , and its  $t$  statistic for the null hypothesis  $\alpha = 1$ , for all values of the truncation lag parameter  $k$  between 1 and 12. In general, the results are quite robust to which value of  $k$  is selected.<sup>10</sup>

The results presented in this section are quite striking. The unit root hypothesis can be rejected for all but three series. To obtain these results, only a rather weak postulate needed to be imposed, namely the presence of a one-time change in the trend function. We claim that this is a weak postulate for the following reasons. As shown by Nelson and Plosser and Campbell and Mankiw, all the series analyzed have a unit root if the trend function is not allowed to change. This view implies that the 1929 crash is simply one big outlier in the innovation sequence. On the other hand, it also implies that the post-1973 growth slowdown is a succession of smaller innovations and that the mean of the innovations is different for the pre-1973 and the post-1973 period. These alternative interpretations are, we think, less appealing than the hypothesis of a break in the trend function, especially given that we allow such a break under both the null and alternative hypotheses.

Given that the "consumer price," "velocity," and "interest rate" series appear to be characterized by the presence of a unit root, it seems worthwhile to see if this feature is stable in the pre and post 1929 samples. The subsample Dickey-Fuller type regressions are presented in Appendix B for values of  $k$  between 1 and 12. Consider first the "consumer price index." With  $k = 2$ , the estimated value of the sum of the autoregressive coefficients,  $\tilde{\alpha}$  is 0.965 for the pre-1929 sample with a  $t$  statistic of  $-1.28$ . We therefore cannot reject the null hypothesis of a unit root for this subperiod. However, for the post-1929 sample, the picture is rather different; when  $k = 7$ ,  $\tilde{\alpha}$  has a value of 0.704 with a  $t$  statistic of  $-4.56$  significant at the 1 percent level. The nonrejection of the unit root hypothesis using the full sample is due to the pre-1929 sample. After 1929, the unit root is no longer present.

The case for the "velocity" series has a special feature. It has been well documented that the U.S. velocity series declined steadily until 1946, remained at a fairly constant level until 1970, before increasing (see, for example, Poole (1988) and Gould and Nelson (1974)). For this reason, we discuss applications of the standard Dickey-Fuller procedure for the following three samples: 1869–1929, 1930–1945, 1946–1970. We cannot reject the unit root hypothesis with the pre-1929 sample even though  $\tilde{\alpha}$  is only 0.865 ( $k = 0$ ). However, the picture is

<sup>9</sup>Basically, the same estimates of the sum of the autoregressive coefficient  $\alpha$  and its  $t$  statistic were obtained using regressions (13) and (14), for all values of  $k$ , showing some robustness for the results presented. We also applied Phillips' nonparametric procedure to the detrended series (equations (6) and (7),  $i = B$ ). These test statistics did not allow the rejection of the null hypothesis of a unit root.

<sup>10</sup>One notable exception is the quarterly real GNP series where the  $t$  statistic is significant at the 5 percent level with  $k = 2$  or  $k = 10$ . It is significant at the 2.5 percent level with  $k = 11$ , with a value of  $-4.32$ . We choose to report the result of  $k = 10$  because the 10th lagged first-difference was highly significant ( $t$  statistic of 2.29) while the 11th and 12th lags were not.

dramatically different using post-1929 samples. For the period 1930–1945, with  $k = 1$ ,  $\tilde{\alpha}$  is estimated at  $-0.011$  with a  $t$  statistic (for testing  $\alpha = 1$ ) of  $-3.44$  which is significant at the 5 percent level. Though this result should be taken with caution due to the small number of observations, it is suggestive of a quite different behavior. The period 1946–1970 affords a modest amount of additional information and yields still more dramatic results: with  $k = 4$ ,  $\tilde{\alpha}$  is equal to  $0.00$  with a  $t$  statistic (for  $\alpha = 1$ ) of  $-5.86$ , significant at the 1 percent level using the Dickey-Fuller critical values. It therefore appears that the “velocity” and “consumer price” series yield similar results: the presence of a unit root before 1929 but not after.

The results for the interest rate series indicate that, in this case, the unit root hypothesis cannot be rejected at usual significance levels for both subsamples (even though  $\alpha$  is estimated at  $0.540$  ( $k = 3$ ) with the pre-1929 sample). Indeed, given our previous results, we can conclude that only the “interest rate” series is characterized by the presence of a unit root after 1929. All the other series are better construed as stationary fluctuations around a deterministic trend function for this period.

#### 6. DISCUSSION AND CONCLUDING COMMENTS

When testing for the presence of a unit root in a time series of data against the hypothesis of stationary fluctuations around a deterministic trend function, the use of a long span of data has definite advantages. It allows tests with larger power compared to using a smaller span, in most cases even if the latter allows more observations (see Shiller and Perron (1985) and Perron (1987)). The drawback, however, is that a data set with a large span has more chance to include a major event which one would rather consider as an outlier or as exogenous given its relative importance. The arguments in this paper rest on the postulate that two such events have occurred in the 20th century: the 1929 Great Crash and the slowdown in growth after the oil shock of 1973. We therefore considered, as a relevant alternative, a trend function with a change in the intercept in 1929 and a change in the slope after 1973.

Let us discuss, in more detail, what are the relevant issues in drawing particular conclusions about the nature of economic fluctuations from our results. It is particularly important to put our results into perspective and also highlight what has not been shown.

The first important issue to point out is that we have not provided a formal unconditional statistical model of the time series properties of the various aggregates. A rejection of the null hypothesis of a unit root conditional on the possibility of shifts in the underlying trend function at known dates does not imply that the various series can be modeled as stationary fluctuations around a completely deterministic breaking trend function. As a matter of general principle, a rejection of the null hypothesis does not imply acceptance of a particular alternative hypothesis. However, since the tests were designed to have power against a specific class of alternative hypotheses, it is useful to look among close members of that class to propose an interesting statistical model for the various

aggregates. Only with such a model is it possible to provide forecasts with appropriate standard errors.

We certainly do not entertain the view that the trend function including its changes are deterministic. This would imply that one would be able to forecast with certainty future changes. This is indeed quite unappealing. What we have in mind in specifying our class of maintained hypotheses can be parameterized as follows:

$$(16) \quad y_t = \eta_t + Z_t, \quad \eta_t = \mu_t + \beta_t t,$$

where  $A(L)Z_t = B(L)e_t$ ;  $e_t \sim \text{i.i.d.}(0, \sigma^2)$ ;  $\mu_t = \mu_{t-1} + V(L)v_t$ , and  $\beta_t = \beta_{t-1} + W(L)w_t$ . Here, the  $Z_t$ 's are (not necessarily stationary) deviations from the trend function  $\eta_t$ . The intercept and the slope of the trend functions,  $\mu_t$  and  $\beta_t$ , are themselves random variables modeled as integrated processes with  $W(L)$ ,  $V(L)$  stationary and invertible polynomials. However, the important distinction is that the timing of the occurrence of the shocks  $v_t$  and  $w_t$  are rare relative to the sequence of innovations  $\{e_t\}$ , for example, poisson processes with arrival rates specified such that their occurrences are rare relative to the frequency of the realizations in the sequence  $\{e_t\}$ . The intuitive idea behind this type of modeling is that the coefficients of the trend function are determined by long-term economic fundamentals (e.g., the structure of the economic organization, population growth, etc.) and that these fundamentals are rarely changed. In our examples,  $v_t$  is nonzero in 1929 (the great depression) and  $w_t$  is nonzero in 1973 (the oil price shock).

In this sense, our exogeneity assumption about the changes in the trend function is a device that allows taking these shocks out of the noise function into the trend function without specific modeling of the stochastic nature of the behavior of  $\mu_t$  and  $\beta_t$ . It is in this sense that our approach does not provide an unconditional representation of the time series properties of the various variables.

Estimation of models of the form (16) by specifying a probability distribution for the error sequences  $\{e_t, w_t, v_t\}$  is clearly an important avenue of future research. Interesting recent advances on this topic have been provided by Hamilton (1987) and Lam (1988) where the slope of the trend function is allowed to take two different values and the changes are modeled as a binomial process. However, no methods are currently available to test whether  $Z_t$  is integrated or not in this framework.<sup>11</sup> Problems in estimation of models of the form (16) are further compounded by the fact that, according to our view, only one nonzero realization of both  $v_t$  and  $w_t$  would be present in the data set typically available for the series of interest.

In the above framework, the purpose of this paper is to test whether  $Z_t$  is an integrated process or not, i.e. to test whether the shocks  $\{e_t\}$  have persistent effects that do not vanish over a long horizon. Our approach is to remove from the noise function two events that occurred at two dates where we believe positive occurrences of the shocks  $\{v_t, w_t\}$  happened and to model them as part

<sup>11</sup>Both authors studied the behavior of the postwar quarterly real GNP with possible shifts in the slope of the trend function. Hamilton (1987) imposes  $Z_t$  to be an integrated process while Lam (1988) leaves  $Z_t$  unconstrained.

of the trend function. The fact that we model these changes as exogenous implies that our results are conditional. That is, conditional upon the presence of a change in the trend function in 1929 and 1973, the fluctuations are transitory (i.e.,  $Z_t$  is stationary).

An important direction for future research is to make this conditional result into an unconditional statement. This could, in principle, be achieved by a direct test for structural change in the trend function. In a sense, our procedure allows such a test, but conditional on a change occurring at a fixed known date. Hence, problems of pre-testing and "data mining" could be raised regarding the role of looking, ex-post, at the data on the choice of the date.<sup>12</sup> Accordingly, what is needed is a test for structural changes in the trend function occurring at unknown dates. The problem, however, is that care must be applied to ensure that the test has an adequate size under both the unit root and trend-stationary hypotheses. No such test is currently available in the literature. We hope to report, in the near future, developments in this area and applications in this context.

However, an important issue of observational equivalence could not even be settled by such a formal test for structural change. Consider, for instance, the following limiting case in the crash model.<sup>13</sup> A trend-stationary model with a break and where the errors have zero variance is observationally equivalent to a unit root model with drift where the errors have a high probability of being zero but are occasionally nonzero and finite. In general, when the variance of the errors is nonzero, the two models will be nearly observationally equivalent with the disturbances in the unit root model having fat tails. We are able to make a distinction in our empirical result through the mixing conditions (see Assumptions 1 and 2) which prohibit fat tailed disturbances. Any formal test for structural change would, presumably, also have to impose some mixing condition prohibiting fat-tailed disturbances, thereby not resolving this issue of near-observational equivalence.

In fact, any test for the presence of a unit root against trend-stationary alternatives is subject to another type of observational equivalence, as recently argued by Cochrane (1987) and Blough (1988). Indeed, in finite samples, any trend-stationary process is nearly observationally equivalent to a unit root process with a strong mean-reversion component, i.e. where the errors have a moving-average component with a root near minus one.<sup>14</sup> The fact that we reject the unit root hypothesis excluding the event of 1929 suggests that if there is a unit root at all the correlation structure of the innovation sequence must exhibit substantial mean reversion.

<sup>12</sup> See, for example, Christiano (1988).

<sup>13</sup> This issue and its following illustration were raised by a referee.

<sup>14</sup> This observational equivalence problem only disappears asymptotically. In other words, in finite samples, any test for a unit root with ARMA errors should have zero power. Formally, the critical values should be determined such that the test has a given fixed size over all possible values of the nuisance parameters (here, the additional correlation in the errors). Given the near observational equivalence, any such test would have zero power by definition. The unit root tests can be rationalized by arguing that we are willing to have the wrong size over some of the parameter space exactly because for all practical purposes it does not matter whether we label a series as trend stationary or difference stationary with a strong mean-reversion component.

To sum up: trend stationary processes with a break are nearly observationally equivalent to unit root processes with strong mean-reversion and a fat-tailed distribution for the error sequence. Whichever view one adopts cannot be decided by data alone. Nevertheless, the picture under any of these views is basically the same: shocks had little, if any, persistence effect over a long horizon. Only those associated with the Great Depression and the oil price shock significantly altered the long run behavior of the series.

While choosing one view over the other is a matter of convenience for interpreting the data, it has profound implications for a multitude of statistical procedures. Indeed, under the unit root view one must ensure the validity of the procedures under fat-tailed disturbances, and at the moment very few are appropriate in a time series context. Hence for all practical purposes, it may be more advantageous to adopt the trend-stationary view with breaks and detrend our series accordingly prior to analyzing the remaining noise.

*Department of Economics, Princeton University, Princeton, NJ 08544, U.S.A.; and Centre de Recherche et Développement en Économique, Université de Montréal, C.P. 6128, Succ. A, Montréal, Canada, H3C – 3J7.*

*Manuscript received October, 1987; final revision received February, 1989.*

APPENDIX A

The conditions imposed by Assumption 1 permit us to use a functional weak convergence result due to Herrndorf (1984). Let  $S_j = \sum_{i=1}^j e_i (S_0 = 0)$  and define the following variable lying in the space  $D[0, 1]$ :

$$X_T(r) = \sigma^{-1} T^{-1/2} S_{\lfloor Tr \rfloor} = \sigma^{-1} T^{-1/2} S_{j-1}, \quad (j-1)/T \leq r < j/T \quad (j = 1, \dots, T)$$

$$X_T(1) = \sigma^{-1} T^{-1/2} S_T.$$

Herrndorf's theorem states that under the conditions of Assumption 1,  $X_T(r) \Rightarrow w(r)$  where  $\Rightarrow$  denotes weak convergence to the associated probability measure and  $w(r)$  is the unit Wiener process defined on  $C[0, 1]$ , the space of all continuous functions on the interval  $[0, 1]$ . Following Phillips (1987) and Phillips and Perron (1988), it is easy to derive the following lemma related to functions of  $S_j$ . These results will be used in proving both Theorems 1 and 2.

LEMMA A.1: *Let  $S_j = \sum_{i=1}^j e_i (S_0 = 0)$  and assume that the innovation sequence  $\{e_i\}$  satisfies the conditions of Assumption 1. Furthermore, let  $T_B = \lambda T$  for all  $T$ ; then as  $T \rightarrow \infty$ :*

- (a)  $T^{-3/2} \sum_1^{T_B} S_j \Rightarrow \sigma \int_0^\lambda w(r) dr;$
- (b)  $T^{-2} \sum_1^{T_B} S_j^2 \Rightarrow \sigma^2 \int_0^\lambda w(r)^2 dr;$
- (c)  $T^{-5/2} \sum_1^{T_B} j S_j \Rightarrow \sigma \int_0^\lambda r w(r) dr;$
- (d)  $T^{-3/2} \sum_1^{T_B} j e_j \Rightarrow \sigma \left( \lambda w(\lambda) - \int_0^\lambda w(r) dr \right);$
- (e)  $T^{-1} \sum_1^T S_{j-1} e_j \Rightarrow (\sigma^2/2) (w(1)^2 - \sigma_e^2/\sigma^2).$

The results in (a) through (d) are simple extensions of those in Phillips (1987) and Phillips and Perron (1988); part (e) is proved in Phillips (1987). The previous results can be recovered by simply letting  $\lambda = 1$  in which case  $T_B = T$ . For example, to prove part (a):

$$T^{-3/2} \sum_1^{T_B} S_j = \sigma \sum_1^{T_B} \int_{(j-1)/T}^{j/T} T^{-1/2} \sigma^{-1} S_{[Tr]} dr = \sigma \int_0^\lambda X_T(r) dr \Rightarrow \sigma \int_0^\lambda w(r) dr$$

using Herrndorf's weak convergence result and the continuous mapping theorem (see, e.g., Billingsley (1968)). The proofs of (b), (c), and (d) are entirely analogous.

**PROOF OF THEOREM 1:** We consider first the most general model (C) in which both  $\beta$  (the slope) and  $\mu$  (the intercept) are allowed to change, i.e.:

$$(A.1) \quad \begin{aligned} y_t &= \mu_1 + \beta_1 t + e_t, & (0 \leq t \leq T_B). \\ y_t &= \mu_2 + (\beta_2 - \beta_1)T_B + \beta_2 t + e_t, & (T_B < t \leq T). \end{aligned}$$

The following lemma provides a convenient representation of the sample moments of  $\{y_t\}$ .

**LEMMA A.2:** Assume that  $\{y_t\}_0^T$  is generated according to (A.1) with the innovation sequence  $\{e_t\}$  satisfying Assumption 1; then,

$$(a) \quad \begin{aligned} \sum_1^T ty_{t-1} &= (1/6)[\beta_1(3\lambda - \lambda^3) + \beta_2(2 - 3\lambda + \lambda^3)]T^3 \\ &\quad + (\frac{1}{2})[\mu_1\lambda^2 + \mu_2(1 - \lambda^2) + \lambda(1 - \lambda)(\beta_1 - \beta_2)]T^2 \\ &\quad + [(3/2)\mu_1\lambda + (3/2)(1 - \lambda)\mu_2 - \mu_2 - (1/3)\beta_1\lambda - (1/3)\beta_2(1 - \lambda)]T \\ &\quad + \left[ T^{-3/2} \sum_1^T te_{t-1} \right] T^{3/2} + o_p(T), \end{aligned}$$

$$(b) \quad \begin{aligned} \sum_1^T y_{t-1} &= (\frac{1}{2})[\beta_1(\lambda^2 - 2\lambda) + \beta_2(1 + 2\lambda - 3\lambda^2)]T^2 \\ &\quad + [\lambda\mu_1 + (1 - \lambda)\mu_2 + (3/2)\beta_1\lambda - (\frac{1}{2})\beta_2(1 + 3\lambda)]T \\ &\quad + \left[ T^{-1/2} \sum_1^T e_{t-1} \right] T^{1/2} - (\mu_2 - \mu_1), \end{aligned}$$

$$(c) \quad \begin{aligned} \sum_1^T y_{t-1}^2 &= (1/3)[\beta_1^2\lambda^2(3 - 2\lambda) + \beta_2^2(1 - \lambda)^3 + 3\beta_1\beta_2\lambda(1 - \lambda)^2]T^3 \\ &\quad + [\mu_1\beta_1\lambda^2 + \mu_2\beta_2(1 - \lambda^2) - (\frac{1}{2})(\beta_1\lambda + \beta_2(1 - \lambda)^2)]T^2 \\ &\quad + 2 \left[ \lambda(\beta_1 - \beta_2)T^{-1/2} \sum_{T_B+1}^T e_t + \beta_1 T^{-3/2} \sum_1^{T_B} te_t - \beta_2 T^{-3/2} \sum_{T_B+1}^T te_t \right] T^{3/2} \\ &\quad + \left[ \lambda\mu_1^2 + (1/6)\beta_1^2\lambda^2 + \mu_1\beta_1\lambda + (1 - \lambda)\mu_2^2 + (1/6)\beta_2^2(1 - \lambda) \right. \\ &\quad \left. + \mu_2\beta_2(1 - \lambda) - 2\mu_2\beta_2 + T^{-1} \sum_1^T e_{t-1}^2 \right] T + o_p(T), \end{aligned}$$

$$\begin{aligned}
 \text{(d)} \quad \sum_1^T y_t y_{t-1} &= (1/3) [\beta_1^2 \lambda^2 (3 - 2\lambda) + \beta_2^2 (1 - \lambda)^3 + 3\beta_1 \beta_2 \lambda (1 - \lambda)^2] T^3 \\
 &+ [\mu_1 \beta_1 \lambda^2 + \mu_2 \beta_2 (1 - \lambda^2)] T^2 \\
 &+ 2 \left[ \beta_1 T^{-3/2} \sum_1^{T_B} t e_{t-1} + \beta_2 T^{-3/2} \sum_{T_B+1}^T t e_{t-1} + \lambda (\beta_1 - \beta_2) T^{-1/2} \sum_{T_B+1}^T e_t \right] T^{3/2} \\
 &+ [\mu_1^2 \lambda + \mu_2^2 (1 - \lambda) - (1/3) (\lambda \beta_1^2 + (1 - \lambda) \beta_2^2) \\
 &\quad + \lambda (\beta_2^2 - \beta_1 \beta_2) + \lambda \beta_1 (\mu_2 + \mu_1) - 2\lambda \mu_2 \beta_2] T + o_p(T).
 \end{aligned}$$

It is straightforward but tedious to show that  $\tilde{\alpha}$  is given by  $\tilde{\alpha} = A/D$  where

$$\begin{aligned}
 \text{(A.2)} \quad A &= (T^2/2 + T/2) \sum y_{t-1} \sum y_t \\
 &- (T^3/3 + T^2/2 + T/6) \sum y_{t-1} \sum y_t - T \sum y_{t-1} \sum t y_t \\
 &+ (T^2/2 + T/2) \sum y_{t-1} \sum t y_t + (T^4/12 - T^2/12) \sum y_t y_{t-1}
 \end{aligned}$$

and

$$\begin{aligned}
 \text{(A.3)} \quad D &= (T^4/12 - T^2/12) \sum y_{t-1}^2 - T (\sum t y_{t-1})^2 \\
 &+ (T^2 + T) \sum t y_{t-1} \sum y_{t-1} - (T^3/3 + T^2/2 + T/6) (\sum y_{t-1})^2
 \end{aligned}$$

where all summations run from 1 to  $T$ . To prove part (b) we derive the limits of  $T^{-7}A$  and  $T^{-7}D$  using Lemma A.2:

$$\begin{aligned}
 T^{-7}D &= (1/12) T^{-3} \sum y_{t-1}^2 - T^{-6} (\sum t y_{t-1})^2 \\
 &+ T^{-3} \sum t y_{t-1} \cdot T^{-2} \sum y_{t-1} - (1/3) T^{-4} (\sum y_{t-1})^2 + o_p(1) \\
 &= (1/36) [\beta_1^2 \lambda^3 + \beta_2^2 (1 - 3\lambda + 2\lambda^3) + 3\beta_2 \beta_1 \lambda (1 - \lambda^2)] \\
 &- (1/36) [\beta_1 \lambda (3 - \lambda^2) + \beta_2 (2 - 3\lambda + \lambda^3)]^2 \\
 &+ (1/12) [\beta_1 \lambda (3 - \lambda^2) + \beta_2 (2 - 3\lambda + \lambda^3)] [\beta_1 \lambda (2 - \lambda) + \beta_2 (1 - \lambda)^2] \\
 &- (1/12) [\beta_1 \lambda (2 - \lambda) + \beta_2 (1 - \lambda)^2]^2 + o_p(1).
 \end{aligned}$$

Simple algebra yields:

$$\text{(A.4)} \quad T^{-7}D \rightarrow (1/36) (-3\lambda^2 + 4\lambda^3 - 3\lambda^4 + 3\lambda^5 - \lambda^6) (\beta_1 - \beta_2)^2.$$

Similarly,

$$\begin{aligned}
 T^{-7}A &= T^{-3} \sum y_{t-1} \cdot T^{-2} \sum y_{t-1} - (1/3) T^{-4} (\sum y_{t-1})^2 - T^{-6} (\sum t y_{t-1})^2 \\
 &+ (1/12) T^{-3} \sum y_t y_{t-1} + o_p(1) \\
 &= (1/12) [\beta_1 \lambda (3 - \lambda^2) + \beta_2 (2 - 3\lambda + \lambda^3)] \cdot [\beta_1 \lambda (2 - \lambda) + \beta_2 (1 - \lambda)^2] \\
 &- (1/12) [\beta_1 \lambda (2 - \lambda) + \beta_2 (1 - \lambda)^2]^2 \\
 &- (1/36) [\beta_1 \lambda (3 - \lambda^2) + \beta_2 (2 - 3\lambda + \lambda^3)]^2 \\
 &+ (1/36) [\beta_1^2 \lambda^3 + \beta_2^2 (1 - 3\lambda + 2\lambda^3) + 3\beta_2 \beta_1 \lambda (1 - \lambda^2)] + o_p(1)
 \end{aligned}$$

and

$$T^{-7}A \rightarrow (1/36)(-3\lambda^2 + 4\lambda^3 - 3\lambda^4 + 3\lambda^5 - \lambda^6)(\beta_1 - \beta_2)^2.$$

Therefore,  $\tilde{\alpha} = T^{-7}A/T^{-7}D \rightarrow 1$  proving part (b, i). To prove part (b, ii), note that  $T(\tilde{\alpha} - 1) = T^{-6}(A - D)/T^{-7}D$ . Simple manipulation yields:

$$\begin{aligned} T^{-6}(A - D) &= (1/12)T^{-2} \sum y_t y_{t-1} - (1/12)T^{-2} \sum y_{t-1}^2 + T^{-5} \sum t y_{t-1} \sum y_{t-1} \\ &\quad + (1/6)T^{-3} y_T \sum y_{t-1} - (\frac{1}{2})T^{-4} y_T \sum t y_{t-1} - (1/12)T^{-4} (\sum y_{t-1})^2 \\ &\quad + o_p(1) \\ &= (1/24) [\beta_1^2 \lambda^2 + \beta_2^2 (1 - \lambda)^2 + 2\beta_1 \beta_2 \lambda (1 - \lambda)] \\ &\quad + (1/12) [\beta_1 \lambda (3 - \lambda^2) + \beta_2 (2 - 3\lambda + \lambda^3)] [\beta_1 \lambda (2 - \lambda) + \beta_2 (1 - \lambda)^2] \\ &\quad + (1/12) [(\beta_1 - \beta_2) \lambda + \beta_2] \cdot [\beta_1 \lambda (2 - \lambda) + \beta_2 (1 - \lambda)^2] \\ &\quad - (1/12) [(\beta_1 - \beta_2) \lambda + \beta_2] [\beta_1 \lambda (3 - \lambda^2) + \beta_2 (2 - 3\lambda + \lambda^3)] \\ &\quad - (1/12) [\beta_1 \lambda (2 - \lambda) + \beta_2 (1 - \lambda)^2]^2 + o_p(1) \end{aligned}$$

and

$$T^{-6}(A - D) \rightarrow (1/24)(\beta_1 - \beta_2)^2 [-\lambda^2 + 4\lambda^3 - 5\lambda^4 + 2\lambda^5].$$

Therefore,

$$T(\tilde{\alpha} - 1) \rightarrow (3/2)[-1 + 4\lambda - 5\lambda^2 + 2\lambda^3] \cdot [-3 + 4\lambda - 3\lambda^2 + 3\lambda^3 - \lambda^4]^{-1}$$

proving part (b, ii).

**PROOF OF PART (a):** The proof of part (a) is straightforward but tedious. The following arguments present the main steps. First, the expressions for the moments of  $\{y_t\}$  in Lemma A.2 can be written as follows:

$$\sum_1^T t y_{t-1} = a_1 T^3 + b_1 T^2 + c_1 T^{3/2} + d_1 T + o_p(T),$$

$$\sum_1^T y_{t-1} = a_2 T^2 + b_2 T + c_2 T^{1/2} + d_2,$$

$$\sum_1^T y_{t-1}^2 = a_3 T^3 + b_3 T^2 + c_3 T^{3/2} + d_3 T + o_p(T),$$

$$\sum_1^T y_t y_{t-1} = a_4 T^3 + b_4 T^2 + c_4 T^{3/2} + d_4 T + o_p(T).$$

The coefficients  $a_i, b_i, c_i$  and  $d_i$  ( $i = 1, \dots, 4$ ) are given by the corresponding expressions in Lemma A.2 where the condition  $\beta_1 = \beta_2 = \beta$  is imposed. From (A.3), the denominator of  $\tilde{\alpha}$  is given by:

$$\begin{aligned} T^{-5}D &= [a_3/12 - a_1^2 + a_1 a_2 - a_2^2/3] T^2 \\ &\quad + [b_3/12 - 2a_1 b_1 + a_1 b_2 + b_1 a_2 + a_1 a_2 - (2/3)a_2 b_2 - a_2^2/2] T \\ &\quad + [c_3/12 - 2a_1 c_1 + a_1 c_2 + c_1 a_2 - (2/3)a_2 c_2] T^{1/2} \\ &\quad + [d_3/12 - a_3/12 - b_1^2 - 2a_1 d_1 + a_1 d_2 + b_1 b_2 \\ &\quad \quad + d_1 a_2 + a_1 b_2 + b_1 a_2 - b_2^2/3 - (2/3)a_2 d_2 - a_2 b_2 - a_2^2/6] + o_p(1). \end{aligned}$$



Now, tedious algebra shows that the  $O(T^2)$ ,  $O(T)$ , and  $O(T^{1/2})$  coefficients all cancel out and the sum of the  $O(1)$  coefficients yields the following result:

$$T^{-5}D \rightarrow (1/12)\sigma_e^2 + (1/12)[\mu_1 - \mu_2]^2[\lambda - 4\lambda^2 + 6\lambda^3 - 3\lambda^4].$$

To analyze the limit of  $T^{-5}A$ , the numerator of  $\tilde{\alpha}$ , we need the expansions for  $\sum y_i$  and  $\sum ty_i$ . These are given by:

$$\begin{aligned} \sum_1^T ty_i &= a_1T^3 + (b_1 + \beta/2)T^2 + c_1T^{3/2} + (1/3)(d_1 + (5/6)\beta + \mu_2)T + o_p(T), \\ \sum_1^T y_i &= a_2T^2 + (b_2 + \beta)T + c_2T^{1/2} + d_2. \end{aligned}$$

Then we can write, using (A.2):

$$\begin{aligned} T^{-5}A &= [a_1a_2 - a_2^2/3 - a_1^2 + a_4/12]T^2 \\ &+ [a_1a_2 + (b_2 + \beta)(a_1/2 - a_2/3) + (b_1 + \beta/2)(a_2/2 - a_1) - a_2^2/2 + b_4/12 \\ &\quad + a_1b_2/2 - a_1b_1 - a_2b_2/3 + b_1a_2/2]T \\ &+ [a_1c_2 + c_1a_2 - (2/3)a_2c_2 - 2a_1c_1 + c_4/12]T^{1/2} \\ &+ [(b_2 + \beta)(b_1/2 + a_1/2 - b_2/3 - a_2/2) \\ &+ (1/3)(d_1 + (5/6)\beta + \mu_2)(a_2/2 - a_1) \\ &+ (b_1 + \beta/2)(b_2/2 + a_2/2 - b_1) + d_1a_2/2 + b_1a_2/2 - a_2d_2/3 - a_2b_2/2 \\ &\quad - a_2^2/6 - d_1a_1 + a_1d_2/2 + d_4/12 + a_4/12]. \end{aligned}$$

Again, tedious algebra shows that the  $O(T^2)$ ,  $O(T)$ , and  $O(T^{1/2})$  terms all cancel out and that the sum of the coefficients remaining yields the following result:

$$T^{-5}A \rightarrow (1/12)\gamma_1 + (1/12)[\mu_1 - \mu_2]^2[\lambda - 4\lambda^2 + 6\lambda^3 - 3\lambda^4].$$

This proves part (a).

**PROOF OF THEOREM 2:** We first note some invariance properties of the estimators  $\tilde{\alpha}^i$  ( $i = A, B, C$ ):  $\tilde{\alpha}^A$  is invariant with respect to  $\mu$  and  $d$ ;  $\tilde{\alpha}^B$  is invariant to  $\mu_1$  and  $\mu_2$ ; and  $\tilde{\alpha}^C$  is invariant to  $d$ ,  $\mu_1$  and  $\mu_2$ . Hence, without loss of generality, we can study the limiting distribution of  $T(\tilde{\alpha}^i - 1)$  and  $t_{\tilde{\alpha}^i}$  under the null hypothesis that the sequence  $\{y_i\}$  is generated according to:

$$(A.5) \quad y_i = y_{i-1} + e_i \quad (i = 1, \dots, T)$$

with the innovation sequence satisfying Assumption 1. It is also straightforward to show that under (A.5),  $\tilde{\alpha}^i$  ( $i = A, B, C$ ) are asymptotically equivalent to the least-squares estimators  $\hat{\alpha}^i$  ( $i = A, B, C$ ) in the following regressions:

$$(A.6) \quad y_i = \hat{\mu}^A + \hat{\theta}^A DU_i + \hat{\beta}^A t + \hat{\alpha}^A y_{i-1} + \hat{e}_i^A,$$

$$(A.7) \quad y_i = \hat{\mu}^B + \hat{\beta}^B t + \hat{\gamma}^B DT_i^* + \hat{\alpha}^B y_{i-1} + \hat{e}_i^B,$$

$$(A.8) \quad y_i = \hat{\mu}^C + \hat{\theta}^C DU_i + \hat{\beta}^C t + \hat{\gamma}^C DT_i + \hat{\alpha}^C y_{i-1} + \hat{e}_i^C,$$

where  $DU_i = 1$  and  $DT_i = DT_i^* = 0$  if  $t \leq T_B$  and  $DU_i = 0$ ,  $DT_i^* = t - T_B$ ,  $DT_i = t$  if  $t > T_B$ . Since we are also concerned with regressions of the type (A.6)–(A.8) later in the text, we shall derive the limiting distributions concerning  $\tilde{\alpha}^i$  ( $i = A, B, C$ ) using the representation of  $\hat{\alpha}^i$  ( $i = A, B, C$ ). Note, however, that  $\hat{\alpha}^A$  in (A.6) and  $\hat{\alpha}^C$  in (A.9) are not invariant to the value of the parameter  $d$  under the null hypothesis. To achieve invariance, one must introduce a dummy variable  $D(TB)_i$  taking value 1 at  $t = T_B + 1$  and 0 elsewhere, as is done in regressions (12) and (14). Under the null hypothesis,  $\hat{\alpha}^B$  is not invariant to a drift taking two distinct values:  $\mu_1$  and  $\mu_2$ . To achieve invariance, the variable  $DU_i$  must be introduced in (A.7). However, this affects the limiting distribution of  $\hat{\alpha}^B$  which becomes equivalent to that of  $\hat{\alpha}^C$ . Hence, one cannot analyze directly the case of a joint segmented trend function in a one-step type regression. We now turn to the proof of the theorem deriving the limiting distribution of  $T(\hat{\alpha}^i - 1)$  and  $t_{\hat{\alpha}^i}$  ( $i = A, B, C$ ) in (A.6) through (A.8).

The following Lemma provides weak convergence results for the sample moments of the data and will be used extensively. Its proof is a simple extension of Lemma A.1 and follows the methods of Phillips (1987) and Phillips and Perron (1988).

LEMMA A.3: Let  $\{y_t\}_0^T$  be a stochastic process generated according to (A.5) with the innovation sequence satisfying Assumption 1. Furthermore, let  $T_B = \lambda T$ ; then as  $T \rightarrow \infty$ :

- (a)  $T^{-3/2} \sum_1^{T_B} y_t \Rightarrow \sigma \int_0^\lambda w(r) dr,$
- (b)  $T^{-5/2} \sum_1^{T_B} t y_t \Rightarrow \sigma \int_0^\lambda r w(r) dr,$
- (c)  $T^{-2} \sum_1^{T_B} y_t^2 \Rightarrow \sigma^2 \int_0^\lambda w(r)^2 dr,$
- (d)  $T^{-1} \sum_1^T y_{t-1} e_t \Rightarrow (\sigma^2/2)(w(1)^2 - \sigma_\varepsilon^2/\sigma^2).$

Parts (a), (b), and (c) are simple generalizations of results in Phillips (1987) and Phillips and Perron (1988) where  $T_B = T$  and hence  $\lambda = 1$ ; part (d) is proved in Phillips (1987).

Using the property that a regression of the form  $y = X_1 \hat{\beta}_1 + X_2 \hat{\beta}_2 + \hat{\varepsilon}$  yields a numerically identical estimator  $\hat{\beta}_2$  as obtained in a regression of the form  $y^* = X_2^* \hat{\beta}_2 + \hat{\varepsilon}$  where  $y^*$  and  $X_2^*$  are projections of  $y$  and  $X_2$ , respectively, on the space spanned by the vectors in  $X_1$ , (A.6), (A.7), and (A.8), can be written as:

$$(A.9) \quad Y_{it} = \hat{\beta}^i X_{1it} + \hat{\psi}_i X_{2it} + \hat{\alpha}^i y_{it-1}^* + \hat{\varepsilon}_{it} \quad (t = 1, \dots, T; i = A, B, C),$$

where

$$\begin{aligned} \hat{\psi}_A &= \hat{\theta}^A; \quad \hat{\psi}_B = \hat{\gamma}^B; \quad \hat{\psi}_C = \hat{\gamma}^C, \\ Y_{At} &= Y_{Bt} = y_t - \bar{Y}, \\ y_{At-1}^* &= y_{Bt-1}^* = y_{t-1} - \bar{Y}_{-1}, \\ X_{1At} &= X_{1Bt} = t - \bar{i}, \\ X_{1Ct} &= t - c_1 \quad \text{if } t \leq T_B \quad \text{and } = t - c_2 \quad \text{otherwise,} \\ X_{2At} &= DU_t - \lambda, \\ X_{2Bt} &= -\bar{i}^* \quad \text{if } t \leq T_B \quad \text{and } = t - T_B - \bar{i}^* \quad \text{otherwise,} \\ X_{2Ct} &= t - c_1 \quad \text{if } t \leq T_B \quad \text{and } = 0 \quad \text{otherwise,} \\ Y_{Ct} &= y_t - A \quad \text{if } t \leq T_B \quad \text{and } = y_t - B \quad \text{otherwise,} \\ y_{Ct-1}^* &= y_{t-1} - A' \quad \text{if } t \leq T_B \quad \text{and } = y_{t-1} - B' \quad \text{otherwise,} \end{aligned}$$

and

$$\begin{aligned} \bar{i} &= T^{-1} \sum_1^T t = (T+1)/2, \quad \bar{i}^* = T^{-1} \sum_1^{T-T_B} t = T(1-\lambda)^2/2 + (1-\lambda)/2, \\ c_1 &= \bar{i} - T(1-\lambda)/2, \quad c_2 = \bar{i} + T\lambda/2, \\ A &= \bar{Y} + \lambda^{-1} T^{-1} \sum_1^{T_B} (y_t - \bar{Y}), \quad B = \bar{Y} - (1-\lambda)^{-1} T^{-1} \sum_1^{T_B} (y_t - \bar{Y}), \\ A' &= \bar{Y}_{-1} + \lambda^{-1} T^{-1} \sum_1^{T_B} (y_{t-1} - \bar{Y}_{-1}), \quad B' = \bar{Y}_{-1} - (1-\lambda)^{-1} T^{-1} \sum_1^{T_B} (y_{t-1} - \bar{Y}_{-1}), \\ \bar{Y} &= T^{-1} \sum_1^T y_t, \quad \bar{Y}_{-1} = T^{-1} \sum_1^T y_{t-1}. \end{aligned}$$

Now, let  $Y_i' = (y_{i,0}^*, \dots, y_{i,T-1}^*)$ ,  $E' = (e_1, \dots, e_T)$ ,  $Z_i' = [X_{1i}, X_{2i}]$ ,  $X_{1i} = (X_{1i,1}, \dots, X_{1i,T})$ ,  $X_{2i} = (X_{2i,1}, \dots, X_{2i,T})$ ; then, under the null hypothesis,

$$(A.10) \quad \hat{\alpha}^i - 1 = (Y_i' Y_i)^{-1} Y_i' E - (Y_i' Y_i)^{-1} Y_i' Z_i [Z_i' (I - P_{Y_i}) Z_i]^{-1} Z_i' [I - P_{Y_i}] E$$

where  $P_{Y_i} = Y_i (Y_i' Y_i)^{-1} Y_i'$ . Furthermore, define the following terms:

$$Z_i' Z_i = \begin{bmatrix} a_i & b_i \\ b_i & c_i \end{bmatrix}, \quad Z_i' Y_i = \begin{bmatrix} H_i \\ J_i \end{bmatrix}, \quad Z_i' E = \begin{bmatrix} K_i \\ L_i \end{bmatrix}.$$

Straightforward manipulation yields the following representation for  $\hat{\alpha}^i - 1$ :

$$(A.11) \quad \hat{\alpha}^i - 1 = \left[ (a_i c_i - b_i^2) Y_i' E - c_i H_i K_i + b_i J_i K_i - a_i J_i L_i + b_i H_i L_i \right] / \left[ (a_i c_i - b_i^2) Y_i' Y_i - a_i J_i^2 - c_i H_i^2 + 2 b_i H_i J_i \right] \equiv E_i / F_i.$$

The quantities involved are defined as follows with their respective limits as  $T \rightarrow \infty$ , obtained using Lemma A.3:

$$\begin{aligned} a_A &= a_B = T^3/12 - T/12, & T^{-3} a_A &= T^{-3} a_B \rightarrow 1/12, \\ b_A &= -(1-\lambda)\lambda T^2/2, & T^{-2} b_A &\rightarrow -(1-\lambda)\lambda/2, \\ c_A &= (1-\lambda)\lambda T, & T^{-1} c_A &\rightarrow (1-\lambda)\lambda, \\ Y_A' Y_A &= Y_B' Y_B = \sum_1^T (y_{t-1} - \bar{Y}_{-1})^2, \\ T^{-2} Y_A' Y_A &= T^{-2} Y_B' Y_B \Rightarrow \sigma^2 \left[ \int_0^1 w(r)^2 dr - \left( \int_0^1 w(r) dr \right)^2 \right], \\ Y_A' E &= Y_B' E = \sum_1^T (y_{t-1} - \bar{Y}_{-1}) e_t, \\ T^{-1} Y_A' E &= T^{-1} Y_B' E \Rightarrow (\sigma^2/2) (w(1)^2 - \sigma_e^2/\sigma^2) - \sigma^2 w(1) \int_0^1 w(r) dr, \\ H_A &= H_B = \sum_1^T t (y_{t-1} - \bar{Y}_{-1}), \\ T^{-5/2} H_A &= T^{-5/2} H_B \Rightarrow \sigma \int_0^1 r w(r) dr - (\sigma/2) \int_0^1 w(r) dr, \\ J_A &= \sum_1^{T_B} (y_{t-1} - \bar{Y}_{-1}), & T^{-3/2} J_A &\Rightarrow \sigma \int_0^\lambda w(r) dr - \sigma \lambda \int_0^1 w(r) dr, \\ K_A &= K_B = \sum_1^T t u_t - \bar{i} \sum_1^T u_t, & T^{-3/2} K_A &= T^{-3/2} K_B \Rightarrow (\sigma/2) w(1) - \sigma \int_0^1 w(r) dr, \\ L_A &= \sum_1^{T_B} (e_t - \bar{E}), & T^{-1/2} L_A &\Rightarrow \sigma w(\lambda) - \sigma \lambda w(1), \\ b_B &= \sum_{T_B+1}^T (t-i)(t-T_B), & T^{-3} b_B &\rightarrow (1-\lambda)^2 (1+2\lambda)/12, \\ c_B &= T_B \bar{i}^* + \sum_1^{T-T_B} (t-i^*)^2, & T^{-3} c_B &\rightarrow (1-\lambda)^3 (1+3\lambda)/12, \\ J_B &= \sum_{T_B+1}^T (t-T_B)(y_{t-1} - \bar{Y}_{-1}), \\ T^{-5/2} J_B &\Rightarrow \sigma \int_\lambda^1 r w(r) dr - \lambda \sigma \int_\lambda^1 w(r) dr - ((1-\lambda)^2/2) \sigma \int_0^1 w(r) dr, \end{aligned}$$

$$\begin{aligned}
 L_B &= \sum_{T_{B+1}}^T (t - T_B)(e_t - \bar{E}), \quad T^{-3/2}L_B \Rightarrow ((1 - \lambda^2)/2)\sigma w(1) - \sigma \int_{\lambda}^1 w(r) dr, \\
 a_C &= \sum_1^{T_B} (t - c_1)^2 + \sum_{T_{B+1}}^T (t - c_2)^2, \quad T^{-3}a_C \rightarrow (1 - \lambda)^2/12 + \lambda^3/12, \\
 b_C &= \sum_1^{T_B} (t - c_1)^2, \quad T^{-3}b_C \rightarrow \lambda^3/12, \\
 c_C &= \sum_1^{T_B} (t - c_1)^2, \quad T^{-3}c_C \rightarrow \lambda^3/12, \\
 Y'_C Y_C &= \sum_1^{T_B} (y_{t-1} - A')^2 + \sum_{T_{B+1}}^T (y_{t-1} - B')^2, \\
 T^{-2}Y'_C Y_C &\Rightarrow \sigma^2 \left[ \int_0^1 w(r)^2 dr - \lambda^{-1} \left( \int_0^{\lambda} w(r) dr \right)^2 - (1 - \lambda)^{-1} \left( \int_{\lambda}^1 w(r) dr \right)^2 \right], \\
 Y'_C E &= \sum_1^{T_B} (y_{t-1} - A') e_t + \sum_{T_{B+1}}^T (y_{t-1} - B') e_t, \\
 T^{-1}Y'_C E &\Rightarrow (\sigma^2/2)(w(1)^2 - \sigma_e^2/\sigma^2) - \lambda^{-1}\sigma^2 w(\lambda) \int_0^{\lambda} w(r) dr \\
 &\quad - \sigma^2(w(1) - w(\lambda))(1 - \lambda)^{-1} \int_{\lambda}^1 w(r) dr, \\
 H_C &= \sum_1^{T_B} (t - c_1)(y_{t-1} - A') + \sum_{T_{B+1}}^T (t - c_2)(y_{t-1} - B'), \\
 T^{-5/2}H_C &\Rightarrow \sigma \left[ \int_0^1 r w(r) dr - \left(\frac{1}{2}\right)(1 + \lambda) \int_0^{\lambda} w(r) dr + \left(\frac{1}{2}\right) \int_{\lambda}^1 w(r) dr \right], \\
 J_C &= \sum_1^{T_B} (t - c_1)(y_{t-1} - A'), \quad T^{-5/2}J_C \Rightarrow \sigma \left[ \int_0^{\lambda} r w(r) dr - (\lambda/2) \int_0^{\lambda} w(r) dr \right], \\
 K_C &= \sum_1^{T_B} (t - c_1) e_t + \sum_{T_{B+1}}^T (t - c_2) e_t, \\
 T^{-3/2}K_C &\Rightarrow \sigma \left[ \left(\frac{1}{2}\right)(1 - \lambda)w(1) + \left(\frac{1}{2}\right)w(\lambda) - \int_0^1 w(r) dr \right], \\
 L_C &= \sum_1^{T_B} (t - c_1) e_t, \quad T^{-3/2}L_C \Rightarrow \sigma \left[ (\lambda/2)w(\lambda) - \int_0^{\lambda} w(r) dr \right].
 \end{aligned}$$

Now, using (A.11), we can write the statistics as follows:

(A.12)  $T(\hat{\alpha}^A - 1) = T^{-5}E_A/T^{-6}F_A,$

(A.13)  $t_{\hat{\alpha}}A = T^{-5}E_A/[\hat{S}_A^2 \cdot T^{-6}F_A \cdot T^{-4}(a_A c_A - b_A^2)]^{1/2},$

(A.14)  $T(\hat{\alpha}' - 1) = T^{-7}E_i/T^{-8}F_i, \quad i = B, C,$

(A.15)  $t_{\hat{\alpha}}i = T^{-7}E_i/[\hat{S}_i^2 \cdot T^{-8}F_i \cdot T^{-6}(a_i c_i - b_i^2)]^{1/2}, \quad i = B, C,$

where  $\hat{S}_i^2 = T^{-1} \sum_{t=1}^T \hat{e}_i^2$ .

The results of Theorem 2 follow taking the limits of the expressions in (A.12) through (A.15) as  $T \rightarrow \infty$ , using the weak convergence results of the relevant moments (given above) and the fact that  $\hat{S}_i^2 (i = A, B, C)$  converges in probability to  $\sigma_e^2$  as  $T \rightarrow \infty$ .

APPENDIX B

EXTENDED SET OF RESULTS FOR TESTS OF A UNIT ROOT USING SPLIT AND FULL SAMPLES

TABLE A1  
TESTS FOR A UNIT ROOT ON PRE-1929 SAMPLES

		Regression: $y_t = \bar{\mu} + \bar{\beta}t + \bar{\alpha}y_{t-1} + \sum_{i=1}^k \bar{\delta}_i \Delta y_{t-i} + \bar{\epsilon}_t$											
		k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=11	k=12
Real GNP	$\bar{\alpha}$	0.44	0.42	0.43	0.32								
	$t_{\bar{\alpha}}$	-2.33	-1.83	-1.43	-1.27								
Nominal GNP	$\bar{\alpha}$	0.60	0.52	0.45	0.44	0.57							
	$t_{\bar{\alpha}}$	-2.14	-2.04	-1.80	-1.12	-0.89							
Real per capita GNP	$\bar{\alpha}$	0.39	0.37	0.37	0.23								
	$t_{\bar{\alpha}}$	-2.44	-1.91	-1.51	-1.36								
Industrial production	$\bar{\alpha}$	0.69	0.72	0.65	0.64	0.69	0.73	0.68	0.71				
	$t_{\bar{\alpha}}$	-3.14	-2.57	-3.00	-2.83	-2.18	-1.78	-1.93	-1.43				
Employment	$\bar{\alpha}$	0.76	0.80	0.80	0.78	0.82	0.74	0.61	0.56				
	$t_{\bar{\alpha}}$	-2.16	-1.65	-1.72	-1.91	-1.43	-2.05	-2.85	-2.63				
GNP deflator	$\bar{\alpha}$	0.84	0.80	0.78	0.77	0.79	0.75	0.74	0.71				
	$t_{\bar{\alpha}}$	-2.44	-2.79	-2.66	-2.54	-2.10	-2.19	-2.05	-2.05				
Consumer prices	$\bar{\alpha}$	0.95	0.97	0.96	0.96	0.96	0.97	0.96	0.96				
	$t_{\bar{\alpha}}$	-1.89	-1.28	-1.37	-1.50	-1.47	-1.29	-1.37	-1.41				
Wages	$\bar{\alpha}$	0.76	0.73	0.68	0.66	0.63	0.54	0.30	0.43				
	$t_{\bar{\alpha}}$	-2.32	-2.22	-2.35	-2.09	-1.90	-2.08	-2.82	-1.67				
Real wages	$\bar{\alpha}$	0.44	0.50	0.49	0.43	0.24	0.28	0.09	0.68				
	$t_{\bar{\alpha}}$	-2.83	-2.23	-2.05	-1.89	-2.21	-1.71	-1.84	-0.63				
Money stock	$\bar{\alpha}$	0.75	0.71	0.55	0.61	0.55	0.46	0.25	-0.04				
	$t_{\bar{\alpha}}$	-2.88	-2.89	-4.20	-2.71	-2.53	-2.54	-3.01	-3.37				
Velocity	$\bar{\alpha}$	0.89	0.90	0.91	0.90	0.93	0.91	0.89	0.90				
	$t_{\bar{\alpha}}$	-1.52	-1.33	-1.22	-1.17	-0.82	-0.93	-1.12	-1.08				
Interest rate	$\bar{\alpha}$	0.73	0.71	0.54	0.54								
	$t_{\bar{\alpha}}$	-1.73	-1.53	-2.06	-1.59								
Common stock prices	$\bar{\alpha}$	0.81	0.86	0.73	0.75	0.76	0.75	0.64	0.68	0.61			
	$t_{\bar{\alpha}}$	-1.95	-1.27	-2.29	-1.83	-1.60	-1.55	-2.09	-1.75	-2.06			
Quarterly real GNP <sup>a</sup>	$\bar{\alpha}$	0.93	0.91	0.92	0.93	0.93	0.92	0.91	0.90	0.90	0.88	0.86	0.89
	$t_{\bar{\alpha}}$	-2.51	-3.02	-2.52	-2.23	-2.18	-2.41	-2.60	-2.83	-2.68	-3.04	-3.40	-2.62

<sup>a</sup>The sample for Quarterly Real GNP is 47:1-73:1.

TABLE A2  
TESTS FOR A UNIT ROOT ON POST-1929 SAMPLES

Regression:  $y_t = \bar{\mu} + \bar{\beta}t + \bar{\alpha}y_{t-1} + \sum_{i=1}^k \bar{\alpha}_i \Delta y_{t-i} + \bar{\varepsilon}_t$

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=11	k=12
Real GNP	$\bar{\alpha}$ 0.72	0.75	0.75	0.78	0.74	0.66	0.56	0.33	0.22	0.18	0.30	0.09
	$t_{\bar{\alpha}}$ -3.48	-3.00	-2.57	-1.96	-2.16	-2.65	-3.03	-5.32	-3.97	-3.19	-2.05	-2.34
Nominal GNP	$\bar{\alpha}$ 0.76	0.87	0.87	0.89	0.86	0.79	0.72	0.63	0.56	0.51	0.41	0.40
	$t_{\bar{\alpha}}$ -3.37	-1.87	-1.60	-1.35	-1.66	-2.40	-2.93	-4.10	-3.39	-3.20	-2.91	-2.24
Real per capita GNP	$\bar{\alpha}$ 0.78	0.79	0.79	0.82	0.78	0.71	0.63	0.43	0.34	0.32	0.46	0.30
	$t_{\bar{\alpha}}$ -3.05	-2.91	-2.55	-1.96	-2.19	-2.64	-3.03	-5.36	-3.91	-3.08	-1.89	-2.13
Industrial production	$\bar{\alpha}$ 0.68	0.69	0.66	0.63	0.71	0.64	0.53	0.23	0.03	-0.11	-0.03	-0.24
	$t_{\bar{\alpha}}$ -2.84	-2.75	-2.65	-2.52	-1.76	-2.13	-2.55	-5.75	-4.38	-3.61	-2.40	-2.33
Employment	$\bar{\alpha}$ 0.78	0.80	0.80	0.83	0.80	0.72	0.67	0.47	0.55	0.42	0.24	0.10
	$t_{\bar{\alpha}}$ -3.14	-2.66	-2.34	-1.80	-2.02	-2.64	-2.70	-5.61	-2.94	-3.24	-3.30	-2.79
GNP deflator	$\bar{\alpha}$ 0.85	0.92	0.91	0.87	0.87	0.83	0.72	0.76	0.75	0.81	0.75	0.68
	$t_{\bar{\alpha}}$ -2.60	-1.32	-1.68	-2.29	-2.25	-2.64	-4.47	-3.08	-2.57	-1.73	-2.14	-2.48
Consumer prices	$\bar{\alpha}$ 0.84	0.91	0.90	0.88	0.85	0.81	0.70	0.68	0.62	0.72	0.67	0.58
	$t_{\bar{\alpha}}$ -2.92	-1.64	-1.76	-2.10	-2.50	-2.72	-4.55	-3.45	-3.21	-1.98	-2.04	-2.39
Wages	$\bar{\alpha}$ 0.78	0.88	0.89	0.88	0.85	0.81	0.77	0.76	0.78	0.75	0.66	0.60
	$t_{\bar{\alpha}}$ -3.37	-1.69	-1.41	-1.46	-1.77	-2.12	-2.27	-3.19	-2.02	-2.48	-3.06	-2.81
Real wages	$\bar{\alpha}$ 0.72	0.71	0.70	0.75	0.68	0.54	0.58	0.36	0.40	0.23	0.14	-0.35
	$t_{\bar{\alpha}}$ -3.33	-3.10	-2.55	-1.84	-2.14	-2.79	-2.05	-3.31	-2.14	-2.34	-1.89	-2.11
Money stock	$\bar{\alpha}$ 0.89	0.93	0.92	0.91	0.89	0.87	0.86	0.79	0.67	0.63	0.59	0.39
	$t_{\bar{\alpha}}$ -3.01	-1.85	-2.03	-2.55	-2.50	-2.59	-2.57	-4.57	-5.93	-3.61	-2.83	-4.05
Velocity	$\bar{\alpha}$ 0.61	0.67	0.63	0.71	0.69	0.66	0.60	0.52	0.49	0.63	0.50	0.36
	$t_{\bar{\alpha}}$ -3.82	-2.70	-2.80	-1.97	-1.95	-1.91	-2.11	-2.33	-2.16	-1.43	-1.82	-2.54
Interest rate	$\bar{\alpha}$ 1.07	0.98	1.01	1.03	1.04	1.06	0.97					
	$t_{\bar{\alpha}}$ 1.11	-0.28	0.14	0.37	0.51	0.60	-0.25					
Common stock prices	$\bar{\alpha}$ 0.65	0.84	0.79	0.79	0.75	0.64	0.58	0.64	0.62	0.66	0.58	0.52
	$t_{\bar{\alpha}}$ -3.48	-1.58	-1.99	-1.89	-2.13	-3.07	-2.96	-1.99	-2.00	-1.63	-1.64	-1.56
Quarterly real GNP <sup>a</sup>	$\bar{\alpha}$ 0.88	0.84	0.85	0.84	0.82	0.80	0.77	0.80	0.73	0.65	0.52	0.52
	$t_{\bar{\alpha}}$ -2.23	-2.74	-2.39	-2.44	-2.48	-2.63	-2.80	-2.19	-2.78	-3.30	-4.03	-3.17

<sup>a</sup>The sample for Quarterly Real GNP is 73:II-86:III.

TABLE A3  
TESTS FOR A UNIT ROOT ON THE FULL SAMPLES

(1) <sup>a</sup>	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=11	k=12	
Real GNP	(12) $\hat{\alpha}$	0.71	0.68	0.66	0.63	0.62	0.55	0.43	0.28	0.19	0.19	0.15	0.13
	$t_{\hat{\alpha}}$	-4.04	-4.06	-3.86	-3.73	-3.48	-3.87	-4.81	-5.03	-4.89	-4.14	-4.16	-4.20
Nominal GNP	(12) $\hat{\alpha}$	0.70	0.69	0.69	0.70	0.66	0.60	0.56	0.47	0.46	0.41	0.23	0.34
	$t_{\hat{\alpha}}$	-4.43	-4.17	-3.87	-3.58	-4.01	-4.70	-4.88	-5.42	-5.18	-5.38	-7.86	-5.98
Real per capita GNP	(12) $\hat{\alpha}$	0.76	0.73	0.72	0.70	0.70	0.64	0.53	0.43	0.39	0.42	0.37	0.34
	$t_{\hat{\alpha}}$	-3.62	-3.58	-3.37	-3.21	-2.94	-3.27	-4.09	-4.08	-3.89	-3.14	-3.23	-3.17
Industrial production	(12) $\hat{\alpha}$	0.67	0.65	0.59	0.56	0.61	0.57	0.48	0.32	0.40	0.37	0.29	0.32
	$t_{\hat{\alpha}}$	-4.63	-4.46	-4.84	-4.69	-3.83	-3.98	-4.65	-5.47	-4.15	-4.05	-4.45	-4.08
Employment	(12) $\hat{\alpha}$	0.78	0.80	0.77	0.76	0.78	0.73	0.67	0.60	0.64	0.62	0.59	0.58
	$t_{\hat{\alpha}}$	-3.77	-3.29	-3.72	-3.78	-3.12	-3.94	-4.51	-4.76	-3.59	-3.58	-3.59	-3.49
GNP deflator	(12) $\hat{\alpha}$	0.84	0.82	0.81	0.78	0.78	0.75	0.74	0.71	0.70	0.69	0.67	0.68
	$t_{\hat{\alpha}}$	-3.79	-3.99	-3.89	-4.16	-4.04	-4.20	-4.10	-4.32	-4.28	-4.34	-4.35	-4.00
Consumer prices	(12) $\hat{\alpha}$	0.97	0.98	0.97	0.97	0.97	0.97	0.97	0.96	0.96	0.96	0.97	0.97
	$t_{\hat{\alpha}}$	-1.87	-1.28	-1.68	-2.19	-2.06	-1.89	-2.01	-2.01	-1.90	-1.93	-1.49	-1.44
Wages	(12) $\hat{\alpha}$	0.77	0.76	0.74	0.73	0.71	0.68	0.62	0.62	0.64	0.63	0.60	0.67
	$t_{\hat{\alpha}}$	-4.31	-4.15	-4.25	-4.05	-4.21	-4.73	-5.41	-4.91	-4.62	-4.63	-4.69	-3.64
Real wages	(14) $\hat{\alpha}$	0.68	0.63	0.57	0.52	0.49	0.47	0.38	0.30	0.29	0.28	0.27	0.29
	$t_{\hat{\alpha}}$	-3.87	-3.97	-4.11	-4.06	-3.89	-3.62	-4.02	-4.28	-4.11	-3.82	-3.64	-3.37
Money stock	(12) $\hat{\alpha}$	0.87	0.87	0.85	0.86	0.84	0.81	0.78	0.76	0.74	0.74	0.72	0.77
	$t_{\hat{\alpha}}$	-3.94	-3.61	-3.86	-3.66	-3.82	-4.29	-4.69	-4.56	-4.36	-4.29	-4.32	-3.18
Velocity	(12) $\hat{\alpha}$	0.93	0.94	0.95	0.96	0.97	0.97	0.96	0.95	0.96	0.96	0.94	0.98
	$t_{\hat{\alpha}}$	-1.82	-1.53	-1.41	-1.11	-0.74	-0.79	-0.94	-1.09	-0.91	-0.89	-1.22	-0.49
Interest rate	(12) $\hat{\alpha}$	1.01	0.98	0.97	0.98	0.98	0.99	1.004	1.01	1.01	1.03	0.99	0.97
	$t_{\hat{\alpha}}$	0.12	-0.45	-0.53	-0.34	-0.31	-0.24	0.07	0.22	0.12	0.46	-0.08	-0.40
Common stock prices	(14) $\hat{\alpha}$	0.72	0.73	0.72	0.74	0.76	0.76	0.75	0.75	0.73	0.67	0.62	0.60
	$t_{\hat{\alpha}}$	-4.87	-4.39	-4.43	-3.91	-3.53	-3.52	-3.51	-3.41	-3.52	-4.19	-4.50	-4.35
Quarterly real GNP	(10) $\hat{\alpha}$	0.92	0.90	0.91	0.90	0.90	0.90	0.89	0.89	0.88	0.86	0.84	0.87
	$t_{\hat{\alpha}}$	-3.33	-3.97	-3.47	-3.18	-3.16	-3.39	-3.51	-3.42	-3.55	-3.98	-4.32	-3.28

<sup>a</sup>The number in column (1) refers to the regression model used to test for a unit root using the full samples. See the discussion in the text.

## REFERENCES

- BILLINGSLEY, P. (1968): *Convergence of Probability Measures*. New York: John Wiley.
- BLANCHARD, O. J., AND L. H. SUMMERS (1986): "Hysteresis and the European Unemployment Problem," in *NBER Macroeconomics Annual 1986*, ed. by Stanley Fisher. Cambridge: M.I.T. Press, 15-77.
- BLOUGH, S. R. (1988): "On the Impossibility of Testing for Unit Roots and Cointegration in Finite Samples," Working Paper No. 211, John Hopkins University.
- BOX, G. E. P., AND G. C. TIAO (1975): "Intervention Analysis with Applications to Economic and Environmental Problems," *Journal of the American Statistical Association*, 70, 70-79.
- CAMPBELL, J. Y., AND N. G. MANKIW (1987): "Permanent and Transitory Components in Macroeconomic Fluctuations," *American Economic Review, Papers and Proceedings*, 77, 111-117.
- (1988): "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, 102, 857-880.
- CHAN, N. H. (1988): "The Parameter Inference for Nearly Nonstationary Time Series," *Journal of the American Statistical Association*, 83, 857-862.
- CHRISTIANO, L. J. (1988): "Searching for Breaks in GNP," N.B.E.R. Working Paper No. 2695.
- COCHRANE, J. H. (1987): "A Critique of the Application of Unit Root Tests," mimeo, University of Chicago.
- (1988): "How Big is the Random Walk in GNP," *Journal of Political Economy*, 96, 893-920.
- DICKEY, D. A. (1976): "Estimation and Hypothesis Testing for Nonstationary Time Series," Unpublished Ph.D. dissertation, Iowa State University, Ames.
- DICKEY, D. A., AND W. A. FULLER (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427-431.
- (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057-1072.
- ENGLE, R. F., AND C. W. J. GRANGER (1987): "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251-276.
- FULLER, W. A. (1976): *Introduction to Statistical Time Series*. New York: John Wiley and Sons.
- GOULD, J. P., AND C. R. NELSON (1974): "The Stochastic Structure of the Velocity of Money," *American Economic Review*, 64, 405-418.
- HALL, R. E. (1978): "Stochastic Implication of the Life Cycle - Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971-987.
- HAMILTON, J. D. (1987): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycles," *Econometrica*, 57, 357-384.
- HERRNDORF, N. (1984): "A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables," *Annals of Probability*, 12, 141-153.
- LAM, P. S. (1988): "The Generalized Hamilton Model: Estimation and Comparison with Other Models of Economic Time Series," mimeo, The Ohio State University.
- NELSON, C. R., AND C. I. PLOSSER (1982): "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 10, 139-162.
- OULIARIS, S., J. Y. PARK, AND P. C. B. PHILLIPS (1988): "Testing for a Unit Root in the Presence of a Maintained Trend," in *Advances in Econometrics and Modelling*, ed. by Baldev Raj. Dordrecht: Kluwer Academic Publishers, 7-28.
- PERRON, P. (1987): "Test Consistency with Varying Sampling Interval," *Cahier de Recherche*, No. 4187, C.R.D.E., Université de Montréal.
- (1988): "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach," *Journal of Economic Dynamics and Control*, 12, 297-332.
- PHILLIPS, P. C. B. (1987): "Time Series Regression with Unit Roots," *Econometrica*, 55, 277-302.
- PHILLIPS, P. C. B., AND S. N. DURLAUF (1986): "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, 53, 473-496.
- PHILLIPS, P. C. B., AND P. PERRON (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335-346.
- POOLE, W. (1988): "Monetary Policy Lessons of Recent Inflation and Disinflation," *Journal of Economic Perspectives*, 2, 73-100.
- RAPPOPORT, P., AND L. REICHLIN (1987): "Segmented Trends and Nonstationary Time Series," mimeo.
- SAID, S. E., AND D. A. DICKEY (1984): "Testing for Unit Roots in Autoregressive - Moving Average Models of Unknown Order," *Biometrika*, 71, 599-608.
- SAMUELSON, P. A. (1973): "Proof that Properly Discounted Present Values of Assets Vibrate Randomly," *The Bell Journal of Economics and Management Science*, 4, 369-374.



- SHILLER, R. J., AND P. PERRON (1985): "Testing the Random Walk Hypothesis: Power Versus Frequency of Observation," *Economics Letters*, 18, 381-386.
- STOCK, J. H., AND M. W. WATSON (1988): "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097-1107.
- STULZ, R. M., AND W. WASSERFALLEN (1985): "Macroeconomic Time Series, Business Cycles and Macroeconomic Policies," in *Understanding Monetary Regimes*, ed. by Karl Brunner and Allan H. Metzler. Carnegie-Rochester Conference Series on Public Policy, 22, 9-54.
- TSAY., R. S. (1986): "Time Series Model Specification in the Presence of Outliers," *Journal of the American Statistical Association*, 81, 132-141.
- WASSERFALLEN, W. (1986): "Non-Stationarities in Macro-Economic Time Series — Further Evidence and Implications," *Canadian Journal of Economics*, 19, 498-510.